

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 06:51:36 ; Search time 120.153 Seconds
(without alignments)
3219.229 Million cell updates/sec

Title: US-09-821-883-4
Perfect score: 3783
Sequence: 1 MRAAPLLIARAASLSGLF.....CWKPVQKAPPPPAHHHHH 697

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DB=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO/spool/US09821883/runat_09092004_105128_6657/app_query.fasta_1.3100
-DB=Issued Patents_NA -QFMT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=Dts -START=1 -END=-1 -MATRIX=Dblum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09821883 @CIGN 1.1 215 @runat_09092004_105128_6657 -NCPU=6 -ICPU=3
-NO MWAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents_NA:
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2439.5	64.5	4530	1	US-08-229-515A-9
2	2439.5	64.5	4530	1	US-08-645-865-9
3	2439.5	64.5	4530	4	US-09-167-322-4
4	2439.5	64.5	4530	4	US-09-527-487-1
5	2439.5	64.5	4530	4	US-09-877-177A-11
6	2434.5	64.4	4473	3	US-09-048-804-1
7	2434.5	64.4	4473	3	US-09-056-105-26
8	2434.5	64.4	4473	4	US-09-663-834A-3
9	2434	64.3	9274	4	US-09-811-115-1
10	2422	64.0	3768	4	US-09-811-115-2
11	2417	63.9	3768	2	US-08-625-101-1
12	2417	63.9	3768	2	US-08-356-786-1

13	2033.5	53.8	2385	2	US-09-146-283-3	Sequence 3, Appl1
14	2033.5	53.8	2385	3	US-08-579-823A-3	Sequence 3, Appl1
15	2033.5	53.8	2385	3	US-09-344-195-3	Sequence 3, Appl1
16	1976	52.2	3955	1	US-08-229-515A-14	Sequence 14, Appl
17	1976	52.2	3955	1	US-08-645-865-14	Sequence 14, Appl
18	1626.5	43.0	1872	3	US-08-422-108-2	Sequence 2, Appl1
19	1626.5	43.0	1872	4	US-08-422-734-2	Sequence 2, Appl1
20	732	19.3	1868	1	US-08-658-883B-1	Sequence 1, Appl1
21	732	19.3	1868	4	US-09-676-610B-26	Sequence 26, Appl
22	731.5	19.3	1593	4	US-09-676-610B-25	Sequence 25, Appl
23	731.5	19.3	5532	2	US-08-475-035-3	Sequence 3, Appl1
24	731.5	19.3	5532	4	US-09-676-610B-17	Sequence 17, Appl
25	723.5	19.1	5484	3	US-09-632-580A-3	Sequence 3, Appl1
26	723.5	19.1	5501	1	US-08-484-438-1	Sequence 3, Appl1
27	723.5	19.1	5555	1	US-08-484-438-3	Sequence 3, Appl1
28	711	18.8	4905	1	US-07-978-895-3	Sequence 3, Appl1
29	711	18.8	4905	1	US-08-473-119-3	Sequence 3, Appl1
30	711	18.8	4905	2	US-08-475-352-3	Sequence 3, Appl1
31	711	18.8	4905	4	US-09-170-659-3	Sequence 3, Appl1
32	711	18.8	4975	3	US-09-630-706-3	Sequence 3, Appl1
33	703	18.6	1958	4	US-09-570-454-1	Sequence 1, Appl1
34	703	18.6	1958	4	US-09-867-521-1	Sequence 1, Appl1
35	624.5	16.5	4545	6	5183884-3	Patent No. 5183884
36	533.5	14.1	1588	2	US-09-146-283-1	Sequence 1, Appl1
37	533.5	14.1	1588	3	US-08-579-823A-1	Sequence 1, Appl1
38	533.5	14.1	1588	3	US-09-344-195-1	Sequence 1, Appl1
39	493	13.0	322	1	US-08-421-356-1	Sequence 1, Appl1
40	493	13.0	322	4	US-09-046-783-1	Sequence 1, Appl1
41	491.5	13.0	1409	4	US-09-634-368-3	Sequence 3, Appl1
42	490.5	13.0	896	6	5200327-2	Patent No. 5200327
43	490.5	13.0	900	1	US-08-318-193-7	Sequence 7, Appl1
44	487	12.9	905	6	5200327-3	Patent No. 5200327
45	487	12.9	909	1	US-08-318-193-9	Sequence 9, Appl1

ALIGNMENTS

RESULT 1
US-08-229-515A-9
; Sequence 9, Application US/08229515A
; Patent No. 5518885
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESS: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,515A
; FILING DATE: 19 APR 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-9880
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-229-515A-9

```

Alignment Scores:

Pred. No.:	2,48-178	Length:	4530
Score:	2439.50	Matches:	562
Percent Similarity:	40.61%	Conservative:	28
Best Local Similarity:	38.68%	Mismatches:	76
Query Match:	64.49%	Indels:	789
DB:	1	Gaps:	15

US-09-821-883-4 (1-697) x US-08-229-515A-9 (1-4530)

Qy	2	ArgAlaAlaProLeuLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly	17
Db	94	CGCGCCCTCCAGCGGGGTCTCAGCGCGGACCATGGGCGGAGCCGCGATGAGCCACCATG	153
Qy	18	PhelLeuPheLeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArg	37
Db	154	GAGCTGGCGCGCTTGTGTCCGCTGG-----GGGCTCCTCTCGCCCTCTTTGCCCC	204
Qy	38	GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer	57
Db	205	GGAGCCCGGACACCCCAAGTGTGCACCGGCACAGACATGAAGCTGGGGTCTCCTGCCAGT	264
Qy	58	ProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGly	77
Db	265	CCCGAGACCCACCTGGACATGCTCCGCCACCTCTACCAAGGCTGCAGGTGGTCAGGGA	324
Qy	78	AsnLeuGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln	97
Db	325	AACCTGGAACCTCACCTCTGCCACCAATGCCAGCTGTCTCTCTGAGGATATCCAG	384
Qy	98	GluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg	117
Db	385	GAGGTGCAGGGCTACGTGCTCATCGCTCACACCAAGTGAAGCAGGTCCCCTGCAGG	444
Qy	118	LeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAsp	137
Db	445	CTCGGATTGTGCGAGGCACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTGTAGAC	504
Qy	138	AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg	157
Db	505	AATGGAGACCGCGTGAAACAATACCACCCCTGTGCAGGGGCTCTCCACAGAGGCTTCGG	564
Qy	158	GluLeuGlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsn	177
Db	565	GAGCTGCAGCTTCGNAGCCTCACAGAGATCTTGAAGGAGGGGTCTTGTATCAGCGGAAC	624
Qy	178	ProGlnLeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGln	197
Db	625	CCCCAGCTCTGCTACCAAGACACGATTTTGTGGAAGGACATCTTCCACAAGAACCAACAG	684
Qy	198	LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet	217
Db	685	CTGGCTCTCACACTGATAGACCAACCGCTCTCGGGCTTCGACCCCTGTCTCCGATG	744
Qy	218	CysLysGlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThr	237
Db	745	TGTAAAGGCTCCCGCTGCTGGGAGAGAGTTCTGAGGATTTGTCAAGGCTGACGGGCAC	804
Qy	238	ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlu	257
Db	805	GTCTGTGCCGCTGGCTGTGCCGCTGCAAGGGGCCACTGCCCACTACTGCTGCCATGAG	864
Qy	258	GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisSph	277
Db	865	CAGTGTGCTGCCGCGTGCACGGGCCCCAAGCACTCTGACTGCGCTGGCCTGCCCTCCACTTC	924
Qy	278	AsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThr	297

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:57:41 ; Search time 4699.88 Seconds
(without alignments)
4428.608 Million cell updates/sec

Title: US-09-821-883-4

Perfect score: 3783

Sequence: 1 MRAAPLLIARAASLSGLF.....CWKPVQKAPPPPAHHHHH 697

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2 1/USPTO.spool/US09821883/runat 09092004 105127 6643/app query.fasta_1.3100
-DB=EST -QFMT=fastap -SUPFIX=rest -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=500 -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09821883 @CGN 1 11571 @runat 09092004 105127 6643 -NCPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: em_esta:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_inv:*
18: em_gss_hum:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vri:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1970.5	52.1	4323	11	AK031099	Mus muscu AK031099 Mus muscu
2	1960	51.8	4463	11	AK083669	Mus muscu AK083669 Mus muscu
3	1291.5	34.1	1016	12	BM562913	AGENCOURT BM562913 AGENCOURT
4	1169	30.9	1129	12	BM802792	AGENCOURT BM802792 AGENCOURT
5	1167.5	30.9	885	14	CA455074	AGENCOURT CA455074 AGENCOURT
6	1166	30.8	871	14	CA488274	AGENCOURT CA488274 AGENCOURT
7	1162	30.7	4715	11	AF318349	Homo sapi AF318349 Homo sapi
8	1139.5	30.1	893	14	CA455141	AGENCOURT CA455141 AGENCOURT
9	1090	28.8	894	14	CA454131	AGENCOURT CA454131 AGENCOURT
10	1090	28.8	1004	14	CD515356	AGENCOURT CD515356 AGENCOURT
11	1089	28.8	653	12	BM721340	UI-E-E01-BM721340 UI-E-E01-
12	1081	28.6	902	14	CA488868	AGENCOURT CA488868 AGENCOURT
13	1073	28.4	588	13	EX478931	DKFZp6866A EX478931 DKFZp6866A
14	1046	27.7	569	9	AL701765	DKFZp6866P AL701765 DKFZp6866P
15	1033	27.3	583	13	BX479114	DKFZp6866P BX479114 DKFZp6866P
16	1022.5	27.0	795	14	CA328613	UI-M-PY0-CA328613 UI-M-PY0-
17	1014	26.8	3110	11	AK031542	Mus muscu AK031542 Mus muscu
18	1010	26.7	808	14	CA489534	AGENCOURT CA489534 AGENCOURT
19	991	26.2	616	14	CD721801	OJ01f07.Y CD721801 OJ01f07.Y
20	967	25.6	943	13	BQ958632	AGENCOURT BQ958632 AGENCOURT
21	966	25.5	1988	11	BC023725	Mus muscu BC023725 Mus muscu
22	966	25.5	3372	11	BC046553	Mus muscu BC046553 Mus muscu
23	935	24.7	855	13	BU594980	AGENCOURT BU594980 AGENCOURT
24	924	24.4	813	12	BI557797	603236977 BI557797 603236977
25	921	24.3	614	10	AW370693	QV1-BT026 AW370693 QV1-BT026
26	909.5	24.0	852	14	CD516283	AGENCOURT CD516283 AGENCOURT
27	908	24.0	791	12	BI154872	602902857 BI154872 602902857
28	903	23.9	717	14	CB598701	AGENCOURT CB598701 AGENCOURT
29	901	23.8	491	12	BM790293	K-RST0070 BM790293 K-RST0070
30	897	23.7	906	14	CA454570	AGENCOURT CA454570 AGENCOURT
31	880	23.3	998	12	BI649877	603296516 BI649877 603296516
32	875	23.1	685	9	AU123871	AU123871 AU123871
33	871	23.0	932	14	CA487981	AGENCOURT CA487981 AGENCOURT
34	868	22.9	609	10	AW701942	UQ93b02.Y AW701942 UQ93b02.Y
35	862.5	22.8	621	12	BG283493	602407782 BG283493 602407782
36	855	22.6	887	13	BQ717097	AGENCOURT BQ717097 AGENCOURT
37	848	22.4	786	12	BI155788	603904360 BI155788 603904360
38	835.5	22.1	649	9	AA496412	zv37a02.r AA496412 zv37a02.r
39	817	21.6	533	14	CB437620	685491.MA CB437620 685491.MA
40	817	21.6	552	12	BG277542	ux45d07.Y BG277542 ux45d07.Y
41	817	21.6	670	9	AI906012	RC-BT105-AI906012 RC-BT105-
42	808	21.4	876	13	BQ769889	UI-M-F10-BQ769889 UI-M-F10-
43	805	21.3	514	10	BF998814	QV0-GN014 BF998814 QV0-GN014
44	800	21.1	887	14	CA980253	AGENCOURT CA980253 AGENCOURT
45	792.5	20.9	678	12	BI555157	603236396 BI555157 603236396

ALIGNMENTS

RESULT 1
AK031099

LOCUS

DEFINITION Mus musculus 13 days embryo forelimb cDNA, RIKEN full-length enriched library, clone:593040N10 product:v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian), full insert sequence.

ACCESSION AK031099

VERSION AK031099.1 GI:26082143

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
10149636

2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3
Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaishiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 4323)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koude, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

FEATURES
source
Location/Qualifiers
1..4323
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

```

```

/db_xref="FANTOM DB:5930404N10"
/db_xref="MGI:2394277"
/db_xref="taxon:10090"
/clone="5930404N10"
/tissue_type="forelimb"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
48..3818
/notes="putative
v-erb-b2 erythroblastic leukemia viral oncogene homolog 2,
neuro/glioblastoma derived oncogene homolog (avian)
(MGI:MGI:954110, GH|U71126, evidence: BLASTN, 99%,
match=449)"

misc_feature
48..3818
/notes="putative
v-erb-b2 erythroblastic leukemia viral oncogene homolog 2,
neuro/glioblastoma derived oncogene homolog (avian)
(MGI:MGI:954110, GH|U71126, evidence: BLASTN, 99%,
match=449)"

ORIGIN

Alignment Scores:
Pred. No.: 1.49e-111 Length: 4323
Score: 1970.50 Matches: 437
Percent Similarity: 36.79% Conservative: 31
Best Local Similarity: 34.36% Mismatches: 73
Query Match: 52.09% Indels: 731
DB: 11 Gaps: 4

US-09-821-883-4 (1-697) x AK031099 (1-4323)

Qy 3 AlaAlaProLeuLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPheLeu 22
Db 33 GCCTGAGTATCATCATGGAGCTGGCGGCTGGTGGGGTTCCTCTCGCCCTC 92
Qy 23 PhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAlaSerThr 42
Db 93 -----CTGTCCCCCGAGCGCGGGTACC 116
Qy 43 GlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThrHisLeu 62
Db 117 CAAGTGTGTACCGGTACCGACATGAAGTTCGCACTCCCTGCCAGTCTCTGAGACCCACTG 176
Qy 63 AspMetLeuArgHisLeuTyrglnGlyCysGlnValValGlnGlnGlyAsnLeuGluLeuThr 82
Db 177 GACATGCTCCGCCACCTCTACCAAGGCTGTTCAGGTGGTCAGGGCAATTTGGAGCTTACC 236
Qy 83 TyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGlnGlyTyr 102
Db 237 TACCTGCCCGCAATGCCAGCTCTCATTCCTGCAGGACATCCAGGAAGTCCAGGGATAC 296
Qy 103 ValLeuIleAlaHisGlnGlnValArgGlnValProLeuGlnArgGlyLeuValArg 122
Db 297 ATGCTCATCGCTCACACCGGAGTGAACACGCTCCCATCTGCAGAGTTTCCGATCTGAGA 356
Qy 123 GlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAspProLeu 142
Db 357 GGGACTCAGCTCTTTGAGGACAAAGTATGCCCTGGCTGTGTAGACAACCGAGACCTTTG 416
Qy 143 AsnAsn---ThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGlnLeu 161
Db 417 GACAACTGACCAACCGCGGCCGCCAGGCAAGAACCCAGAGGGCTGCGGGAGTGCAGCTT 476
Qy 162 ArgSerLeuThrGluIleLeuLysGlyValLeuLeuGlnArgAsnProGlnLeuCys 181
Db 477 CGAAGTCTCAGAGATCTTTGAAGGGGAGGAGTTTTCCTCGTGGGAACCTTCAGCTCTGC 536
Qy 182 TyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeuThr 201
Db 537 TACAGGACATGGTTTGTGGAAGGATGCTCCCGTGAAGATAACACAGTGGCTCTCTGTC 596
Qy 202 LeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGlySer 221
Db 597 GACATGGACCAATCGTTCCTCCGGGCTGTCCACTTGTGTCCTCCCAACCTGCAAGACAAT 656
Qy 222 ArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrValCysAlaGly 241
Db 657 CACTGTGGGGTGAGATCTCTGAAGACTGTGAGATCTTGTGTCACCATCTGTACTAGT 716

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 15:10:16 ; Search time 842.741 Seconds
(without alignments)
4161.385 Million cell updates/sec

Title: US-09-821-883-4
Perfect score: 3783
Sequence: 1 MRAAPLLARAASLSGLF.....CWKPVQKAPPPPAHHHHH 697

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3304383 seqs, 2515761380 residues

Total number of hits satisfying chosen parameters: 6608766

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09821883/runat_09092004_105129_6736/app_query.fasta_1.3100
-DB=Published Applications NA -QWITS=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09821883 @CGN 1.1 1949 @runat_09092004_105129_6736
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
------------	-------	-------	--------	----	-------------

ALIGNMENTS

RESULT 1

US-09-821-883-9
; Sequence 9, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Gradis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2091
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500*-rGM-CSF construct

1	3783	100.0	2091	9	US-09-821-883-9	Sequence 9, Appl
2	3473.5	91.8	2070	9 <td>US-09-821-883-7</td> <td>Sequence 7, Appl</td>	US-09-821-883-7	Sequence 7, Appl
3	3014	79.7	1692	9 <td>US-09-821-883-8</td> <td>Sequence 8, Appl</td>	US-09-821-883-8	Sequence 8, Appl
4	2957.5	78.2	1665	9 <td>US-09-821-883-6</td> <td>Sequence 6, Appl</td>	US-09-821-883-6	Sequence 6, Appl
5	2478	65.5	1437	9 <td>US-09-821-883-10</td> <td>Sequence 10, Appl</td>	US-09-821-883-10	Sequence 10, Appl
6	2442.5	64.6	4543	9 <td>US-09-769-508-1</td> <td>Sequence 1, Appl</td>	US-09-769-508-1	Sequence 1, Appl
7	2439.5	64.5	4530	9 <td>US-09-877-177-11</td> <td>Sequence 11, Appl</td>	US-09-877-177-11	Sequence 11, Appl
8	2439.5	64.5	4530	13 <td>US-10-392-113-45</td> <td>Sequence 45, Appl</td>	US-10-392-113-45	Sequence 45, Appl
9	2439.5	64.5	4530	13 <td>US-10-426-836-11</td> <td>Sequence 11, Appl</td>	US-10-426-836-11	Sequence 11, Appl
10	2439.5	64.5	4530	15 <td>US-10-177-293-125</td> <td>Sequence 125, Appl</td>	US-10-177-293-125	Sequence 125, Appl
11	2439.5	64.5	4530	15 <td>US-10-007-926A-119</td> <td>Sequence 119, Appl</td>	US-10-007-926A-119	Sequence 119, Appl
12	2439.5	64.5	4530	15 <td>US-10-338-730-1</td> <td>Sequence 1, Appl</td>	US-10-338-730-1	Sequence 1, Appl
13	2439.5	64.5	4530	15 <td>US-10-101-510-124</td> <td>Sequence 124, Appl</td>	US-10-101-510-124	Sequence 124, Appl
14	2439.5	64.5	4530	16 <td>US-10-116-275-131</td> <td>Sequence 131, Appl</td>	US-10-116-275-131	Sequence 131, Appl
15	2439.5	64.5	4530	16 <td>US-10-272-437A-27</td> <td>Sequence 27, Appl</td>	US-10-272-437A-27	Sequence 27, Appl
16	2439.5	64.5	4530	16 <td>US-10-117-937-595</td> <td>Sequence 595, Appl</td>	US-10-117-937-595	Sequence 595, Appl
17	2439.5	64.5	4530	16 <td>US-10-159-563-208</td> <td>Sequence 208, Appl</td>	US-10-159-563-208	Sequence 208, Appl
18	2439.5	64.5	4530	16 <td>US-10-435-896-10</td> <td>Sequence 10, Appl</td>	US-10-435-896-10	Sequence 10, Appl
19	2439.5	64.5	4530	17 <td>US-10-734-564-59</td> <td>Sequence 59, Appl</td>	US-10-734-564-59	Sequence 59, Appl
20	2439.5	64.5	4642	15 <td>US-10-198-846-10896</td> <td>Sequence 10896, A</td>	US-10-198-846-10896	Sequence 10896, A
21	2434.5	64.4	4473	10 <td>US-09-441-411-5</td> <td>Sequence 5, Appl</td>	US-09-441-411-5	Sequence 5, Appl
22	2434.5	64.4	4473	15 <td>US-10-146-473-32</td> <td>Sequence 32, Appl</td>	US-10-146-473-32	Sequence 32, Appl
23	2434.5	64.4	4473	15 <td>US-10-207-555-44</td> <td>Sequence 44, Appl</td>	US-10-207-555-44	Sequence 44, Appl
24	2434.5	64.4	4473	15 <td>US-10-101-510-81</td> <td>Sequence 81, Appl</td>	US-10-101-510-81	Sequence 81, Appl
25	2434	64.3	9274	9 <td>US-09-811-123-7</td> <td>Sequence 7, Appl</td>	US-09-811-123-7	Sequence 7, Appl
26	2434	64.3	9274	9 <td>US-09-811-115-1</td> <td>Sequence 1, Appl</td>	US-09-811-115-1	Sequence 1, Appl
27	2422	64.0	3765	15 <td>US-10-207-498-5</td> <td>Sequence 5, Appl</td>	US-10-207-498-5	Sequence 5, Appl
28	2422	64.0	3768	9 <td>US-09-811-123-8</td> <td>Sequence 8, Appl</td>	US-09-811-123-8	Sequence 8, Appl
29	2422	64.0	3768	9 <td>US-09-811-115-2</td> <td>Sequence 2, Appl</td>	US-09-811-115-2	Sequence 2, Appl
30	2422	64.0	3768	13 <td>US-09-984-092-3</td> <td>Sequence 3, Appl</td>	US-09-984-092-3	Sequence 3, Appl
31	2422	64.0	3768	13 <td>US-10-280-576-3</td> <td>Sequence 3, Appl</td>	US-10-280-576-3	Sequence 3, Appl
32	2422	64.0	3768	17 <td>US-10-441-779C-3</td> <td>Sequence 3, Appl</td>	US-10-441-779C-3	Sequence 3, Appl
33	2417	63.9	3768	9 <td>US-09-854-356-9</td> <td>Sequence 9, Appl</td>	US-09-854-356-9	Sequence 9, Appl
34	2417	63.9	3768	9 <td>US-09-930-125-1</td> <td>Sequence 1, Appl</td>	US-09-930-125-1	Sequence 1, Appl
35	2417	63.9	3768	13 <td>US-09-785-973-1</td> <td>Sequence 1, Appl</td>	US-09-785-973-1	Sequence 1, Appl
36	2417	63.9	3768	15 <td>US-10-313-644-1</td> <td>Sequence 1, Appl</td>	US-10-313-644-1	Sequence 1, Appl
37	2414.5	63.8	4606	10 <td>US-09-971-392-70</td> <td>Sequence 70, Appl</td>	US-09-971-392-70	Sequence 70, Appl
38	1976	52.2	3955	9 <td>US-09-870-759-117</td> <td>Sequence 117, Appl</td>	US-09-870-759-117	Sequence 117, Appl
39	1976	52.2	3955	9 <td>US-09-854-356-10</td> <td>Sequence 10, Appl</td>	US-09-854-356-10	Sequence 10, Appl
40	1976	52.2	3955	10 <td>US-09-751-708A-117</td> <td>Sequence 117, Appl</td>	US-09-751-708A-117	Sequence 117, Appl
41	1974.5	52.2	3771	9 <td>US-09-854-356-11</td> <td>Sequence 11, Appl</td>	US-09-854-356-11	Sequence 11, Appl
42	1644	43.5	2132	16 <td>US-10-412-804A-3</td> <td>Sequence 3, Appl</td>	US-10-412-804A-3	Sequence 3, Appl
43	1641.5	43.4	2149	16 <td>US-10-412-804A-9</td> <td>Sequence 9, Appl</td>	US-10-412-804A-9	Sequence 9, Appl
44	1638.5	43.3	2164	16 <td>US-10-412-804A-5</td> <td>Sequence 5, Appl</td>	US-10-412-804A-5	Sequence 5, Appl
45	1587	42.0	867	9 <td>US-09-821-883-24</td> <td>Sequence 24, Appl</td>	US-09-821-883-24	Sequence 24, Appl

US-09-821-883-9

Alignment Scores:

Pred. No.: 0 Length: 2091
Score: 3783.00 Matches: 697
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-821-883-4 (1-697) x US-09-821-883-9 (1-2091)

Qy 1 MetArgAlaProLeuLeuLeuAlaArgAlaSerLeuLeuGlyPheLeuPhe 20
Db 1 ATGAGAGCTGACCCCTCTCTGGCCAGGCGAGCAAGCCTTAGGCTTTGGCTTTCTGTTT 60
Qy 21 LeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyValaAla 40
Db 61 CTGCTTTTCTGGCTAGACCGAGTGTACTAGCCAGGAGTTGGCGCGGGCGCGCG 120
Qy 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuLeuProAlaSerProGluThr 60
Db 121 TCGACCCCAAGTGTGACCGGCACACATGAGCTGGCGCTCCCTGGCCAGTCCCGAGACC 180
Qy 61 HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyValLeuGlu 80
Db 181 CACCTGGACATGCTCCGCCACCTCTACACAGGCTGCCAGGTGGTGCAGGGAACCTGGAA 240
Qy 81 LeuThrTyrLeuProThrAsnAlaSerLeuLeuSerPheLeuGlnAspIleGlnGluValGln 100
Db 241 CTCACCTACCTGCCACCAATGCCAGCTCTCTCTCTGCCAGGATATCCAGGAGGTGCAG 300
Qy 101 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
Db 301 GGCTACGTGCTCATCGCTCACAACTGAGGAGGTCCTCCACTGCAGAGGCTGGCGATT 360
Qy 121 ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 140
Db 361 GTGCGAGGACCCAGCTCTTTGAGGACCACTATGCTCCCTGGCGCTGTAGCAATGGAGAC 420
Qy 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln 160
Db 421 CCGCTGAACATACCAACCTGTACAGGGGCTCCCGAGGAGGCTGGGGAGGTGCAG 480
Qy 161 LeuArgSerLeuThrGluLeuLeuLysGlyValLeuIleGlnArgAsnProGlnLeu 180
Db 481 CTTTGAAGCCTCACAGAGATCTTGAAGGAGGGTCTTGATCCAGCGGAACCCCGAGCTC 540
Qy 181 CysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAlaLeu 200
Db 541 TGCTACCGAGGACACGATTTTGGGAAGGACATCTTCCCAAGAACCAACAGCTGGCTCTC 600
Qy 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220
Db 601 ACATGATAGACACCAACCGCTCTCGGGCTGCGACCCCTGTCTCCGATGTGTAAAGGC 660
Qy 221 SerArgCysTrpGlyLysSerGluAspCysGlnSerLeuThrArgThrValCysAla 240
Db 661 TCCCGCTGTGGGGAGAGATTCTGAGGATTGTGAGAGCTTGACGCGCACTGTCTGTGCC 720
Qy 241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCysAla 260
Db 721 GGTGGCTGTGGCGGTGCAAGGGGCGCCTGCGCCACTGACTGTGTCATGAGCAGGTGTGCT 780
Qy 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
Db 781 GCCGCTGCACGGGCCCCAGACACTCTGACTGCTGGCTGGCTGCCATTCACACACAGT 840
Qy 281 GlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSer 300
Db 841 GGCACTCTGTAGCTGCACTGCCCGAGCCCTGGTCACCTACCAACACACAGACAGCTTGGTCC 900
Qy 301 MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr 320

Db 901 ATGCCCAATCCGAGGCGCGGTATACATTGGCGCCAGCTGTGTACTGCTGTCCCTAC 960
Qy 321 AsnTyrLeuSerThrAspValGlySerAlaSerIleIleAsnPheGluLysLeuGlyAla 340
Db 961 AACTACCTTTCTACGGACGTGGGATCCGCTAGCATCATTAATTTGGAAGAGTTGGGCCCT 1020
Qy 341 GlyGlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeu 360
Db 1021 GGGGGCATGTGTCACACAGGACCGGAGCTCATCTACACGAGTGGGGTGGGACCTG 1080
Qy 361 ThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGlu 380
Db 1081 ACCTAGGGCTGGAGCCCTCTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Qy 381 GlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGln 400
Db 1141 GGGGCTGGCTCCGATGATTTGATGGTGTACCTGGGAATGGGGGAGGAGGAGGAGGAGGAG 1200
Qy 401 SerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValPro 420
Db 1201 AGCTTCCCAACATGACCCCGGCTCTACAGGGTACAGTGAAGGAGGAGGAGGAGGAGGAG 1260
Qy 421 LeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyr 440
Db 1261 CTGCGCTCTGAGACTGATGGCTACGTTGCCCGCCCTGACCTGCAGCGCCCGAGCTGAATAT 1320
Qy 441 ValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAla 460
Db 1321 GTGAACACGAGCAGATGTTCCGCGCCCGAGCCCTTTCGCCCGAGAGGAGGAGGAGGAGGAG 1380
Qy 461 AlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsnGly 480
Db 1381 GCCGACCTGCTGGTGGCCACTCTGGAAGGGCCAGACTCTCTCCCGAGGAGAGAGTGGG 1440
Qy 481 ValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrPro 500
Db 1441 GTCGTCAAGACGTTTTCCTTTGGGGGTGCGGTGGAGAACCCCGAGTACTTGACACC 1500
Qy 501 GlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsn 520
Db 1501 CAGGAGGAGCTGCCCTTCAGCCCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1560
Qy 521 LeuTyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGly 540
Db 1561 CTCTATTACTGGGACCGAGACCCACAGAGGGGGGGCTCCACCCAGGACCTTCAAGGG 1620
Qy 541 ThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaAlaProThr 560
Db 1621 ACACCTACCGCAGAGAACCCAGAGTACCTGGGTCTGGAGCTGCCAGCGCGCCGCCACC 1680
Qy 561 ArgSerProAsnProValThrArgProTrpLysHisValAspAlaIleLysGluAlaLeu 580
Db 1681 CGCTCACCAACCTGTCTCCCGGGCTTGGAGGACATGTAGATGCCATCAAGAGAGCTCTG 1740
Qy 581 SerLeuLeuAsnAspMetArgAlaLeuGluAsnGluLysAsnGluAspValAspIleIle 600
Db 1741 AGCTCTCTAAATGACATGCGTCTCTGGAGAACGAAAGAACGAAAGACGTAAGCATCATC 1800
Qy 601 SerAsnGluPheSerIleGlnArgProThrCysValGlnThrArgLeuLysLeuTyrLys 620
Db 1801 TCTAATGATTTCTCATCCAGAGGCGGACATGTGTGGAGAGCCCGCTGAAAGCTATACAAG 1860
Qy 621 GlnGlyLeuArgGlyAsnLeuThrLysLeuAsnGlyAlaLeuThrMetIleAlaSerHis 640
Db 1861 CAGGCTCTACGGGGCAACCTCACCAAACTCAATGGCGCTTGACCATGATAGCCAGCCAC 1920
Qy 641 TyrGlnThrAsnCysProProThrProGluThrAspCysGluIleGluValThrThrPhe 660
Db 1921 TACCAGAGCACTGCGCTCCAAACCCCGGAACCTGACTGTGAATAGAGTCAACCACTTT 1980
Qy 661 GluAspPheIleLysAsnLeuLysGlyPheLeuPheAspIleProPheAspCysTrpLys 680

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:04:16 ; Search time 686.431 Seconds
(without alignments)
4313.608 Million cell updates/sec

Title: US-09-821-883-4

Perfect score: 3783

Sequence: 1 MRAAPLLAASLSGLFLF.....CWKPVQKGAPPPPAHHHHH 697

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cpn2.1/USPTO.spool/US09821883/runat_09092004.105126.6623/app_query.fasta_1.3100
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09821883 @CNC 1.1 1646 @runat_09092004.105126.6623 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:.*
1: geneseqn1980s:.*
2: geneseqn1990s:.*
3: geneseqn2000s:.*
4: geneseqn2001as:.*
5: geneseqn2001bs:.*
6: geneseqn2002s:.*
7: geneseqn2003as:.*
8: geneseqn2003bs:.*
9: geneseqn2003cs:.*
10: geneseqn2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3783	100.0	2091	AAD21567	Aad21567 Human HER
2	3473.5	91.8	2070	AAD21565	Aad21565 Human HER
3	3014	79.7	1692	AAD21566	Aad21566 Human HER
4	2957.5	78.2	1665	AAD21564	Aad21564 Human HER
5	2478	65.5	1437	AAD21568	Aad21568 Human HER
6	2439.5	64.5	4530	AAT01585	Aat01585 Her-2/neu
7	2439.5	64.5	4530	AAT71253	Aat71253 Human HER
8	2439.5	64.5	4530	AAZ60815	Aaz60815 Nucleotid

9	2439.5	64.5	4530	4	AAD19731	Aad19731 Human tyz
10	2439.5	64.5	4530	6	ABN85585	Abn85585 Human HER
11	2439.5	64.5	4530	6	ABZ35012	Abz35012 Human gen
12	2439.5	64.5	4530	6	ABV94128	Abv94128 Breast ca
13	2439.5	64.5	4530	6	ABK83918	Abk83918 Human CDN
14	2439.5	64.5	4530	7	ACC50139	Acc50139 Breast ca
15	2439.5	64.5	4530	7	ABQ83856	Abq83856 Human Her
16	2439.5	64.5	4530	8	AAD58073	Aad58073 Human C-e
17	2439.5	64.5	4530	9	ADC09594	Adc09594 Her2/Neu
18	2439.5	64.4	4473	2	ABQ76220	Abq76220 Human tum
19	2434.5	64.4	4473	2	AAZ31071	Aaz31071 HER-2 nuc
20	2434.5	64.4	4473	6	ABZ34969	Abz34969 Human gen
21	2434.5	64.4	4473	6	AAD38904	Aad38904 Human Her
22	2434.5	64.4	4473	7	ACC69999	Acc69999 Human C-e
23	2434.5	64.4	4473	9	ADC35148	Adc35148 Human bre
24	2434.5	64.4	4473	9	ADD25483	Add25483 Binding d
25	2434	64.3	9274	6	ABK14057	Abk14057 Human HER
26	2434	64.3	9274	6	ABK14057	Abk14057 Human HER
27	2431.5	64.3	4472	3	AAA14812	Aaa14812 CDNA enco
28	2427	64.2	9274	4	AAF24297	Aaf24297 HER2 tran
29	2422	64.0	3765	9	ADB67620	Adb67620 Human epi
30	2422	64.0	3768	3	AAA09455	Aaa09455 Human her
31	2422	64.0	3768	6	ABV78168	Abv78168 Human ERB
32	2422	64.0	3768	6	ABZ35744	Abz35744 Human ERB
33	2422	64.0	3768	6	AAD43935	Aad43935 Human HER
34	2422	64.0	3768	6	ABX09987	Abx09987 Human ERB
35	2422	64.0	3768	6	AAD43986	Aad43986 Human HER
36	2422	64.0	3768	6	ABK14058	Abk14058 Human HER
37	2422	64.0	3768	6	ABL91709	Ab191709 Human pol
38	2422	64.0	3768	7	ACC57649	Acc57649 Human pro
39	2417	63.9	3768	2	AAT40739	Aat40739 HER-2/neu
40	2417	63.9	3768	2	AAH01912	Aah01912 Human HER
41	2417	63.9	3768	5	AAH23392	Aah23392 Human HER
42	2417	63.9	3768	6	AAD32743	Aad32743 Human Her
43	2417	63.9	3768	6	ABK10730	Abk10730 Human Her
44	2417	63.9	3768	6	ABA92350	Ab92350 Human Her
45	2414.5	63.8	4606	9	ADB47370	Adb47370 Human CDN

ALIGNMENTS

RESULT 1

AAD21567
ID AAD21567 standard; DNA; 2091 BP.

XX AC AAD21567;

XX DT 28-JAN-2002 (first entry)

XX DE Human HER500-rGM-CSF fusion DNA construct comprising OVA-derived peptide.

XX KW Immunostimulatory fusion protein; IFP; antigen component; therapy;
KW Immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
KW HER-2 protein; granulocyte-macrophage colony stimulating factor;
KW ovalbumin-derived octapeptide; OVA; rat; HER500-rGM-CSF fusion DNA; da.

XX OS Homo sapiens.

OS Rattus norvegicus.

OS Unidentified.

OS Chimeric.

XX Key Location/Qualifiers

XX CDS 1..2091

FT /product= a

FT FT comprising human PAP signal sequence, mature PAP protein,
FT FT an Ala Arg linker, human HER-2 signal sequence, mature
FT FT HER-2 membrane distal extracellular domain, an Ala
FT FT linker, an ovalbumin (OVA)-derived immunodominant
FT FT octapeptide, HER-2 membrane distal intracellular domain,

FT an Ala Ala linker, a mature rat GM-CSF sequence and a C-
FT terminal tag"
FT /note= "CDS does not include stop codon"
FT /partial

7.

WO200174855-A2.

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

11-OCT-2001

4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 5

30-MAR-2001.

1007-111-00

30-MAR-2000.

30-MAY-2000;

(DEND-) DENDREON CORP.
 Laus R, Vidovic D, Graddis T;
 WPI: 2001-662965/76.
 P-PSDB; AAE13111.
 An immunostimulatory fusion protein comprising the intracellular domain of HPR-2 and an antigen elicits an immune response to the antigen and is useful for the treatment of associated cancer associated.

.....

Example 1; Page 29; 59pp; English.

Sequence 2091 BP: 447 A; 699 C; 559 G; 386 T; 0 U; 0 Other;

Alignment Scores:

ALPHANUMERIC RECORD:	Pred	No :	Length:	2091
	3	34E-167		

Score:	3783	00
Matches:	3783	00

SCORES:	2703:00
Percent similarity:	100.00%
Conservative:	100.00%
Maximizes:	100.00%

Percent Primality: 100.00%
Percent Local Similarity: 100.00%
Mismatches: 0

BEST LOCAL SIMILARITY:	100.00%	Mismatches:
Overall Match:	100.00%	Indels:

Query Match:	100.00%
Index:	A
Case:	1

US-09-821-883-4 (1-697) x AAD21567 (1-2091)

[illegible][illegible]

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

DB I AIGAGAGCTGCAACCCICCTCCAGGCCAGGGCAAGGCCTAGCCTTGGCTTCCTGTAT 60

[illegible]

QY 21 LeuLeuphefneirypLeuAspargserValLeuAlaIalysGluLeuAlaArgGlyAlaAla 40

100

D6 61 CTGCTTTTTCGGCTAGACCGAAGTGTACTAGCCAAAGAGTTGGCGCGGGGCGG 12

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Qy 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:41:11 ; Search time 7467.3 Seconds
(without alignments)
4045.651 Million cell updates/sec

Title: US-09-821-883-4
Perfect score: 3783
Sequence: 1 MRAAPLLARASLSGLFLF.....CWKPVQKGAPPPPAHHHHH 697

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-Q/cn2_1/USPTO.spool/US09821883/runat_09092004.105127.6631/app_query.fasta_1.3100
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09821883 @CN 1 1 16795 @runat_09092004.105127.6631 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.ev.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.man.*
37: em.htg.vrt.*
38: em.ev.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	3783	100.0	2091	6	AX268288 Sequence
2	3473.5	91.8	2070	6	AX268286 Sequence
3	3014	79.7	1692	6	AX268287 Sequence
4	2957.5	78.2	1665	6	AX268285 Sequence
5	2478	65.5	1437	6	AX268289 Sequence
6	2439.5	64.5	4530	6	I21124 Sequence 9
7	2439.5	64.5	4530	6	I59745 Sequence 9
8	2439.5	64.5	4530	6	AR202597 Sequence
9	2439.5	64.5	4530	6	AR283481 Sequence
10	2439.5	64.5	4530	6	AR344811 Sequence
11	2439.5	64.5	4530	6	AX282577 Sequence
12	2439.5	64.5	4530	6	AX587649 Sequence
13	2439.5	64.5	4530	6	AX644071 Sequence
14	2439.5	64.5	4530	6	AX771418 Sequence
15	2439.5	64.5	4530	6	BD005474 Cellular
16	2439.5	64.5	4530	9	HUMHER2A M11730 Human tyros
17	2434.5	64.4	4473	6	AR080259 Sequence
18	2434.5	64.4	4473	6	AR167390 Sequence
19	2434.5	64.4	4473	6	AR320288 Sequence
20	2434.5	64.4	4473	9	HSERB2R X03363 Human c-erb
21	2434	64.3	9274	6	AR409602 Sequence
22	2434	64.3	9274	6	AX060703 Sequence
23	2422	64.0	3768	6	AR409603 Sequence
24	2422	64.0	3768	6	AX060704 Sequence
25	2422	64.0	3768	6	AX467229 Sequence
26	2422	64.0	3768	6	AX481438 Sequence
27	2422	64.0	3768	6	BD224136 Novel met
28	2417	63.9	3768	6	AR034479 Sequence
29	2417	63.9	3768	6	BD267514 HER-2/neu
30	2417	63.9	3768	6	AX201817 Sequence
31	2417	63.9	3768	6	AX380923 Sequence
32	2417	63.9	3768	6	AX384604 Sequence
33	2417	63.9	3768	6	AX465456 Sequence
34	2346	62.0	3678	6	AX505114 Sequence
35	2142.5	56.6	2763	6	AX380942 Sequence
36	2142.5	56.6	2763	6	AX380944 Sequence
37	2136	56.5	3780	4	AB008451 Canis fam
38	2033.5	53.8	2385	6	AR082744 Sequence
39	2033.5	53.8	2385	6	AR099963 Sequence
40	2033.5	53.8	2385	6	AR143949 Sequence
41	1983.5	52.4	4506	10	AK129487 Mus muscu
42	1976.5	52.2	4727	10	AK116182 Rattus no
43	1976.5	52.2	3955	6	BD267515 HER-2/neu
44	1976	52.2	3955	6	I21129 Sequence 14
45	1976	52.2	3955	6	I59750 Sequence 14

ALIGNMENTS

RESULT 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:28:07 ; Search time 69.255 Seconds
(without alignments)
3175.456 Million cell updates/sec

Title: US-09-821-883-4

Perfect score: 3783

Sequence: 1 MRAAPLLARAASLSGLF.....CWKPVQKAPPPPAHHHHH 697

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_xvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2136	56.5	1259	6 O18735	O18735 canis fam1
2	1962.5	51.9	1259	11 Q8K3P9	Q8K3P9 rattus norv
3	1612.5	42.6	419	4 Q8UKY9	Q8UKY9 homo sapien
4	1380	36.5	711	11 Q80Y89	Q80Y89 mus musculu
5	1014	26.8	881	11 Q8C057	Q8C057 mus musculu
6	966	25.5	367	11 Q8R2X1	Q8R2X1 mus musculu
7	906	23.9	165	4 Q14256	Q14256 homo sapien
8	894.5	23.6	412	4 Q8WYV0	Q8WYV0 homo sapien
9	859	22.7	431	13 Q7SY19	Q7SY19 brachydanio
10	751	19.9	527	13 Q90836	Q90836 gallus gall
11	723	19.1	149	6 Q8EG66	Q8EG66 oryctolagus
12	717.5	19.0	643	11 Q9ERV6	Q9ERV6 mus musculu
13	717.5	19.0	655	11 Q9WVF5	Q9WVF5 mus musculu
14	717.5	19.0	1210	11 Q9EP98	Q9EP98 mus musculu
15	713.5	18.9	1209	11 Q9QX70	Q9QX70 rattus norv
16	708.5	18.7	1209	6 Q8MIL8	Q8MIL8 sus scrofa

17	703	18.6	1191	13 Q7SZP7	Q7SZP7 brachydanio
18	701.5	18.5	478	11 Q9ESE0	Q9ESE0 rattus norv
19	697	18.4	331	4 Q9BUD7	Q9BUD7 homo sapien
20	660	17.4	1165	13 Q9YH40	Q9YH40 xiphophorus
21	658	17.4	1305	13 Q8AW81	Q8AW81 brachydanio
22	635.5	16.8	599	13 Q9PSH2	Q9PSH2 gallus gall
23	610	16.1	1328	13 Q79754	Q79754 fugu rubrip
24	573	15.1	141	11 Q8VH40	Q8VH40 sigmodon hi
25	559.5	14.8	1377	5 Q8MLW0	Q8MLW0 drosophila
26	543	14.4	1433	5 Q9BIH9	Q9BIH9 anopheles g
27	540.5	14.3	1322	5 Q86N22	Q86N22 drosophila
28	471.5	12.5	145	11 Q8CPB5	Q8CPB5 meriones un
29	457	12.1	144	6 Q9MYK4	Q9MYK4 ovine aries
30	448	11.8	146	6 Q8WN17	Q8WN17 equus caball
31	429.5	11.4	152	6 Q9SL10	Q9SL10 equus caball
32	425	11.2	144	6 Q9GL44	Q9GL44 macaca mula
33	417	11.0	141	6 Q7YRF7	Q7YRF7 felis silve
34	411	10.9	144	6 Q86SV5	Q86SV5 papio anubi
35	409.5	10.8	1137	13 Q9W6F6	Q9W6F6 gallus gall
36	400.5	10.6	150	6 Q9BG64	Q9BG64 oryctolagus
37	395	10.4	151	6 Q9BG65	Q9BG65 oryctolagus
38	377	10.0	1368	5 Q23821	Q23821 caenorhabdi
39	341	9.0	366	5 Q26569	Q26569 schistosoma
40	341	9.0	1717	5 Q26566	Q26566 schistosoma
41	331	8.7	334	5 Q26567	Q26567 schistosoma
42	331	8.7	342	5 Q26568	Q26568 schistosoma
43	321.5	8.5	1193	5 Q9Y1X8	Q9Y1X8 ephydatia f
44	308	8.1	138	11 Q99J91	Q99J91 marmota mon
45	305	8.1	1564	5 Q86MD7	Q86MD7 echinococcu

ALIGNMENTS

RESULT 1

Q18735	PRELIMINARY;	PRT; 1259 AA.
AC O18735;		
DT 01-JAN-1998	(Tremblrel. 05, Created)	
DT 01-JAN-1998	(Tremblrel. 05, Last sequence update)	
DT 01-OCT-2003	(Tremblrel. 25, Last annotation update)	
DE ErBB-2.		
OS Canis familiaris (Dog).		
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.		
OX NCBI_taxID=9615;		
RN [1]		
RP SEQUENCE FROM N.A.		
RA Yokota H.;		
RT "CDNA cloning of erbB-2 from canine mammary gland.";		
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.		
DR EMBL; AB008451; BAA23127.1; -		
DR HSSP; P11362; 1FGK.		
DR GO; GO:0016020; C:membrane; IEA.		
DR GO; GO:0005524; F:ATP binding; IEA.		
DR GO; GO:0005509; F:calcium ion binding; IEA.		
DR GO; GO:0005006; F:epidermal growth factor activity; IEA.		
DR GO; GO:0016740; F:transferase activity; IEA.		
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.		
DR GO; GO:0007189; P:transmembrane receptor protein tyrosine kin. . .; IEA.		
DR InterPro; IPR000494; EGFR_L_domain.		
DR InterPro; IPR006211; Furin-like.		
DR InterPro; IPR006212; Furin repeat.		
DR InterPro; IPR009030; Grow_Fac_recep.		
DR InterPro; IPR000719; Prot_kinase.		
DR InterPro; IPR001245; Tyr_kinase.		
DR InterPro; IPR008266; Tyr_kinase_AS.		
DR InterPro; IPR004019; YLP_motif.		
DR Pfam; PF00757; Furin-like; 1.		
DR Pfam; PF00069; pkinase; 1.		
DR Pfam; PF01030; Recep_L_domain; 2.		
DR Pfam; PF02757; YLP; 2.		


```
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00261; FU; 3.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; kinase; transferase; Tyrosine-protein kinase.
SQ SEQUENCE 1259 AA; 137989 MW; E37364D49C4ACD46 CRC64;

Query Match      56.5%; Score 2136; DB 6; Length 1259;
Beet Local Similarity 36.9%; Pred. No. 1.3e-152;
Matches 462; Conservative 22; Mismatches 47; Indels 722; Gaps 3;

Qy 25 WLDRLSVLAKELARGAASQTGCTDMKRLRLPASPTHLDMLRHLYQGQVQVQGNLELYL 84
Db 6 WCRWGLLLALLPSGAAGTQVCTGTDMLKRLRLPASPTHLDMLRHLYQGQVQVQGNLELYL 65

Qy 85 PTNASLFLQDIQEVQGVYLTAHNVQVPLORLRIVRGTQLFEDNYALAVLDNGDPLNN 144
Db 66 PANASLFLQDIQEVQGVYLAHNSQVRQIPQLRLRIVRGTQLFEDNYALAVLDNGDPLEG 125

Qy 145 TTPVTGASPGRLRELQLRLSLTEILKGGVLIQRPOLCYQDTILWKDIFHKNNQALALTID 204
Db 126 GIPAPGAQQGLRELQLRLSLTEILKGGVLIQRPOLCHQDTILWKDIFHKNNQALALTID 185

Qy 205 TNRSRACHPCSPMCKGRWCSESDCSLRTTVCAAGCARCKGLPTDCHEGCAAGCT 264
Db 186 TNRFSACPPSPACKDAHCWAGSGDCQSLRTTVCAAGCARCKGPQPTDCHEGCAAGCT 245

Qy 265 GPKISDCLACLFHNSGICELHCPALVTYNTDTFESMPNPSGRVTFGASCVTACPNYLS 324
Db 246 GPKISDCLACLFHNSGICELHCPALVTYNTDTFESMPNPSGRVTFGASCVTSPNYLS 305

Qy 325 TDVGSASTI----- 333
Db 306 TDVGSCTILVCLNNQVTAEDGTORCEKSKPCARVCYGLGMEHLREVRVTSANIQEFA 365

Qy 334 ----- 333
Db 366 GKCKIFGSLAPESPFDGDPASNTAPLOEQLRVFEALEITGYLIYSAMPDSLPLNSVF 425

Qy 334 ----- 333
Db 426 QNLVRIGRVLHDGAYSLTLQGLGSLWGLSLRELGLAIHRNARLCFVHTVPWDQL 485

Qy 334 ----- 333
Db 486 FRNPHQALLHSANRPEBECVEGLACYCAHGHCHWGPCTQCVNCSQFLRGQECVEECRV 545

Qy 334 ----- 333
Db 546 LQGLPREYVKDRCYCLPCHSECPQNGSVTCFGEADQCVACAHYKDPFPCVARCPGSKVP 605

Qy 334 ----- 333
Db 606 DLSFMP1WKFADEBGTQCPCPINCHSCADLDEKGPAPORASPVTSIIAAVVGILLAV 665

Qy 334 ----- 333
Db 666 VGLVLGILIKRRQKIRKYTMRLLOETELVEPLTPSGAMPNQAMRILKETELRKVKVL 725

Qy 334 ----- 333
Db 726 GSGAGFTVYKGIWIPDGENVKIPVAIKVLRNTPSKANKIILDEAYVMAGVSPVSRLL 785

Qy 334 ----- 333
Db 786 GICLTSTVOLVTQMPYCCLLDHVRHGRGLGSQDLNWCQIAKGSYLEEDVRLVHRDL 845

Qy 334 ----- 333
Db 846 AARNVLKSPNHVKITDFGLARLLIDIDETEHADGGKVPKIKWMALESIPRRRTHQSDVM 905

Qy 334 ----- 333
Db 906 SYGVTVWELMTFGAKPYDGPAREIPDLLEKGERLPOPPICTIDVYIMVKWMIDSECR 965

Qy 334 -NFEKL----- 338
Db 966 PRRELVAESRMARDPQRFVVIQNEDLGPASPLDSTFYRSLLEDGMDGLVDAEYLVP 1025

Qy 339 -----GAGGMVHHRHSSTRSGGDDLTGLGLEPSEEEAPRPLAPSEGAGSDVF 387
Db 1026 QOGFFCPPEPTGAGGTAAHRRHSSTRNGGCELTGLLEPSEEPKPSPLAPSEGAGSDVF 1085

Qy 388 DGLDGMGAAGKLOSLPHDPSPLQRYSEDTVPLPSETDGVVAPLTCSPQPEYVNOQPDVR 447
Db 1086 DGLDGMGAAGKLOSLPQSPSPQRYSEDTVPLPSETDGVVAPLTCSPQPEYVNOQPEVM 1145

Qy 448 PQPSPREGPLPAARPAAGATLER-----AKTLPSPGKGVVVDVFAFGGAVENPEYLTPOG 502
Db 1146 PQPPLALEGLPSPRPAGATLERPKTILSPKTLSPGKGVVVDVFAFGGAVENPEYLAARG 1205

Qy 503 GAAPQHPHPPAFSPAFDNLYWDQDPPERGAAPPSTFKGTPTAENPEYLGLDVP 555
Db 1206 RAAQPHPPHPPAFSPAFDNLYWDQDPSERGSPPSTFEGTPTAENPEYLGLDVP 1258

RESULT 2
Q8K3F9
ID Q8K3F9 PRELIMINARY; PRT; 1259 AA.
AC Q8K3F9;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Neu protooncoprotein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BDIX;
RA Watson P.A., Kim K., Chen K.-S., Gould M.N.;
RT "Androgen-Dependent Mammary Carcinogenesis in Rats Transgenic for the
  Neu Proto-Oncogene.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY116182; AAM5093.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; IEA.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR000494; EGFR_L_domain.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR009030; Grow_fac_recep.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR InterPro; IPR004019; YFP_motif.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; kinase; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR Pfam; PF02757; YLP; 2.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00018; EF_HAND; 1.
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:20:54 ; Search time 12.7941 Seconds
(without alignments)
2836.691 Million cell updates/sec

Title: US-09-821-883-4

Perfect score: 3783

Sequence: 1 MRAAPLLARASLSGLFLF.....CWKPVQKAPPPPAHHHHH 697

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2417	63.9	1255	1 ERB2_HUMAN	P04626 homo sapien
2	1959	51.8	1257	1 ERB2_RAT	P04944 rattus norv
3	1946.5	51.5	1254	1 ERB2_MESAU	Q00553 mesocricetu
4	752	19.9	703	1 EGFR_CHICK	P13387 gallus gall
5	723.5	19.1	1308	1 ERB4_HUMAN	Q15303 homo sapien
6	719.5	19.0	1210	1 EGFR_HUMAN	P00533 homo sapien
7	719.5	19.0	1308	1 ERB4_RAT	Q02956 rattus norv
8	717.5	19.0	1210	1 EGFR_MOUSE	Q01279 mus musculu
9	697	18.4	1342	1 ERB3_HUMAN	P21860 homo sapien
10	683	18.1	127	1 CSF2_RAT	P48750 rattus norv
11	668	17.7	1339	1 ERB3_RAT	Q62799 rattus norv
12	651	17.2	1167	1 YMRK_XIPMA	P13388 xiphophorus
13	539.5	14.3	1426	1 EGFR_DROME	P04412 drosophila
14	466	12.3	141	1 CSF2_MOUSE	P01587 mus musculu
15	461	12.2	144	1 CSF2_SHEEP	P28773 ovis aries
16	444	11.7	144	1 CSF2_CEREL	P51748 cervus elap
17	442	11.7	144	1 CSF2_HUMAN	P04141 homo sapien
18	421.5	11.1	143	1 CSF2_BOVIN	P11052 bos taurus
19	417	11.0	144	1 CSF2_PIG	Q29118 sus scrofa
20	404.5	10.7	1367	1 LRT3_CAEAL	P24348 caenorhabdi
21	396.5	10.5	144	1 CSF2_CANFA	P48749 canis famil
22	388	10.3	144	1 CSF2_FELCA	Q62757 felis silve
23	366.5	9.7	140	1 CSF2_CAVPO	Q60481 cavia porce
24	318	8.4	1363	1 ILPR_BRALA	O02466 branchiost
25	301.5	8.0	245	1 ERB2_MOUSE	P70424 mus musculu
26	276.5	7.3	1477	1 HTK7_HYDAT	Q25197 hydra atten
27	271	7.2	2146	1 INSR_DROME	P09208 drosophila
28	258	6.8	1382	1 INSR_HUMAN	P06213 homo sapien
29	252.5	6.7	1372	1 INSR_MOUSE	P15208 mus musculu
30	252	6.7	1607	1 INPR_LYMST	Q25410 lymanea sta
31	251.5	6.6	1300	1 IRR_MOUSE	Q9wt14 mus musculu
32	249	6.6	1383	1 INSR_RAT	P15127 rattus norv
33	248.5	6.6	1297	1 IRR_HUMAN	P14616 homo sapien

34	246.5	6.5	1300	1 IRR_CAVPO	P14617 cavia porce
35	238.5	6.3	1367	1 IG1R_HUMAN	P08069 homo sapien
36	236.5	6.3	1370	1 IG1R_RAT	P24062 rattus norv
37	234.5	6.2	1373	1 IG1R_MOUSE	Q60751 mus musculu
38	228.5	6.0	581	1 IRR_RAT	Q64716 rattus norv
39	207	5.5	1390	1 INSR_AEDAE	Q93105 aedes aegypt
40	168	4.4	386	1 PPAP_HUMAN	P15309 homo sapien
41	165	4.4	5262	1 MLL2_HUMAN	O14686 homo sapien
42	164	4.3	707	1 SPQ_HUMAN	P23246 homo sapien
43	155.5	4.1	497	1 WAS2_HUMAN	Q9y6w5 homo sapien
44	154.5	4.1	1321	1 IRS2_MOUSE	P81122 mus musculu
45	150	4.0	2715	1 MLL4_HUMAN	Q9umne6 homo sapien

ALIGNMENTS

RESULT 1
ERB2_HUMAN
ID ERB2_HUMAN STANDARD; PRT; 1255 AA.
AC P04626;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell
DE surface receptor HER2) (MLN 19).
GN ERBB2 OR HER2 OR NGL OR NEU.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN 1)
RP SEQUENCE FROM N.A.
RX MEDLINE=86118663; PubMed=3003577;
RA Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N.,
RA Saito T., Toyoshima K.;
RT "Similarity of protein encoded by the human c-erb-B-2 gene to
RT epidermal growth factor receptor.";
RL Nature 319:230-234(1986).
RN 2)
RP SEQUENCE FROM N.A., AND VARIANT ALA-1170.
RX MEDLINE=86070181; PubMed=2999974;
RA Coussens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A.,
RA McGrath J., Seeburg P.H., Libermann T.A., Schlessinger J.,
RA Francke U., Levinson A., Ullrich A.;
RT "Tyrosine kinase receptor with extensive homology to EGF receptor
RT shares chromosomal location with neu oncogene.";
RL Science 230:1132-1139(1985).
RN 3)
RP SEQUENCE FROM N.A., AND VARIANTS CYS-452; VAL-655 AND ALA-1170.
RX Rieder M.J., Livingston R.J., Daniele M.R., Montoya M.A., Chung M.-W.,
RX Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RX Schachwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RN 4)
RP SEQUENCE OF 737-1031 FROM N.A.
RX MEDLINE=86016729; PubMed=2995967;
RA Semba K., Kanata N., Toyoshima K., Yamamoto T.;
RT "A v-erbB-related protooncogene, c-erbB-2, is distinct from the
RT c-erbB-1/epidermal growth factor-receptor gene and is amplified in a
RT human salivary gland adenocarcinoma.";
Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).
RN 5)
RP VARIANTS VAL-654 AND VAL-655.
RX MEDLINE=93194196; PubMed=8095488;
RA Ehsani A., Low J., Wallace R.B., Wu A.M.;
RT "Characterization of a new allele of the human ERBB2 gene by allele-
RT specific competition hybridization.";
Genomics 15:426-429(1993).
RL
CC -1- FUNCTION: Essential component of a neuregulin-receptor complex,
CC although neuregulins do not interact with it alone. GP10 is a
CC potential ligand for this receptor. Not activated by EGF, TGF-

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:29:10 ; Search time 22.2506 Seconds
(without alignments)
3013.200 Million cell updates/sec

Title: US-09-821-883-4

Perfect score: 3783

Sequence: 1 MRAAPLLARASLSGLF.....CWKPVQKAGPPPPAHHHHH 697

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_78.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2417	63.9	1255	1 A24571	protein-tyrosine k
2	1962	51.9	1260	1 TVRTNU	protein-tyrosine k
3	1946.5	51.5	1254	2 I48161	p-185 precursor -
4	752	19.9	1223	1 TVCHLV	epidermal growth f
5	751	19.9	527	2 A42032	epidermal growth f
6	723.5	19.1	1308	2 A47253	epidermal growth f
7	719.5	19.0	1210	1 GQHUE	epidermal growth f
8	717.5	19.0	1210	2 A53183	epidermal growth f
9	713.5	18.9	644	2 A36325	epidermal growth f
10	697	18.4	1342	2 A36223	kinase-related tra
11	683	18.1	127	2 I46293	granulocyte-macrop
12	662	17.5	1339	2 J4387	epidermal growth f
13	651	17.2	1166	1 S06142	protein-tyrosine k
14	539.5	14.3	843	2 A27131	epidermal growth f
15	473	12.5	153	1 F0MSGM	granulocyte-macrop
16	461	12.2	144	2 JH0469	granulocyte-macrop
17	457	12.1	144	1 A61632	granulocyte-macrop
18	442	11.7	144	1 F0HUGM	granulocyte-macrop
19	421.5	11.1	143	1 F0BQGM	granulocyte-macrop
20	404.5	10.7	1323	2 E88257	protein let-23 (im
21	404.5	10.7	1374	2 S70712	protein-tyrosine k
22	396.5	10.5	144	2 A44936	granulocyte-macrop
23	390.5	10.3	1369	2 S70713	protein-tyrosine k
24	379	10.0	1330	1 G0FFE	epidermal growth f
25	341	9.0	366	2 D45558	epidermal growth f
26	341	9.0	1717	1 A45558	epidermal growth f
27	331	8.7	333	2 B45558	epidermal growth f
28	331	8.7	342	2 C45558	epidermal growth f
29	318	8.4	1363	2 T43220	insulin-like growth

ALIGNMENTS

RESULT 1

A24571

protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human

N;Alternate names: c-erb-B-2 protein precursor; kinase-related transforming protein erbB

C;Species: Homo sapiens (man)

C;Date: 25-Oct-1987 #sequence revision 06-Dec-1996 #text change 11-Jun-1999

C;Accession: A24571; A25491; A44188; B44188; I59509; I57622

R;Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T.;

Nature 319, 230-234, 1986

A;Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth f

A;Reference number: A24571; MUID:86118663; PMID:3003577

A;Accession: A24571

A;Molecule type: mRNA

A;Residues: 1-1255 <YAM>

A;Cross-references: GB:X03363; NID:g31197; PIDN:CAA27060.1; PID:g31198

R;Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.

Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985

A;Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epiderm

A;Reference number: A25491; MUID:86016729; PMID:2995967

A;Accession: A25491

A;Molecule type: DNA

A;Residues: 737-1031 <SEM>

A;Cross-references: GB:M11767; NID:g182163; PIDN:AAA35808.1; PID:g553282

R;Coussens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg, P.

Science 230, 1132-1139, 1985

A;Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chromos

A;Reference number: A44188; MUID:86070181; PMID:2999974

A;Accession: A44188

A;Molecule type: DNA

A;Residues: 740-910 <COU1>

A;Cross-references: GB:M12036; NID:g183988; PIDN:AAA35978.1; PID:g183989

A;Accession: B44188

A;Molecule type: mRNA

A;Residues: 1-517; RALL, 522, 'S', 524-654; 'V', 656-1169; 'A', 1171-1255 <COU2>

A;Cross-references: GB:M11730; NID:g183986

R;King, C.R.; Kraus, M.H.; Aaronson, S.A.

Science 229, 974-976, 1985

A;Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.

A;Reference number: I59509; MUID:85272597; PMID:2992089

A;Accession: I59509

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 832-909 <REX>

A;Cross-references: GB:I29395; NID:g459807; PIDN:AAA35809.1; PID:g459808

R;Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.

Mol. Cell. Biol. 7, 2597-2601, 1987

A;Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptional

A;Reference number: I57622; MUID:87286898; PMID:3039351

A;Accession: I57622

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-191 <TAL>

A;Cross-references: GB:ML6792; NID:G183983; PIDN:AAA58637.1; PID:G553332
C;Comment: Amplification and overexpression of this erbB-related gene occurs in about 30
C;Genetics:
A;Gene: GDB:ERBB2; NGL; NEU; HER-2
A;Cross-references: GDB:120613; OMIM:164870
A;Map position: 17q21.1-17q21.1
A;Introns: 25/1; 75/3; 147/1; 883/3
A;Note: the list of introns is incomplete
C;Function:
A;Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
C;Superfamily: epidermal growth factor receptor; protein kinase homology
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
inase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-1255/Product: protein-tyrosine kinase erbB2 #status predicted <MAT>
F;22-653/Domain: extracellular #status predicted <EXT>
F;70-304/Domain: EGF receptor extracellular domain repeat <EE1>
F;395-605/Domain: EGF receptor extracellular domain repeat <EE2>
F;654-675/Domain: transmembrane #status predicted <TM>
F;676-1255/Domain: intracellular #status predicted <INT>
F;718-983/Domain: protein kinase homology <KIN>
F;726-734/Region: protein kinase ATP-binding motif
F;68,124,187,259,530,571,629/Binding site: carbohydrate (Asn) (covalent) #status predict
F;686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F;753/Active site: Lys #status predicted
F;1139,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 63.9%; Score 2417; DB 1; Length 1255;
Best Local Similarity 41.2%; Pred. No. 5.6e-129;
Matches 511; Conservative 4; Mismatches 6; Indels 718; Gaps 2;

Qy	35	LARGAASVCTGTDKMLRIPASPTHLDMLRHLYQGVVQGNLELYLPTNASLFLQ	94
Db	16	LPPGAASVCTGTDKMLRIPASPTHLDMLRHLYQGVVQGNLELYLPTNASLFLQ	75
Qy	95	DIQVQGVYLAHNOVROVPLRLRVRGTQQLFEDNALVALDNGDPLNNTPVTGASPG	154
Db	76	DIQVQGVYLAHNOVROVPLRLRVRGTQQLFEDNALVALDNGDPLNNTPVTGASPG	135
Qy	155	GLRELRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFPHKNQALALTLIDNRRACHPC	214
Db	136	GLRELRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFPHKNQALALTLIDNRRACHPC	195
Qy	215	SPMCKGSRGWGESSEDCSLTRTVCCAGCARGKPLPTDCCEOCACAGCTGPKHSDCLAC	274
Db	196	SPMCKGSRGWGESSEDCSLTRTVCCAGCARGKPLPTDCCEOCACAGCTGPKHSDCLAC	255
Qy	275	LHFNHSGICELHCPALVTYNTDTFESMPNPGRYTFGASCVTACPYNYLSTDVGSASII	333
Db	256	LHFNHSGICELHCPALVTYNTDTFESMPNPGRYTFGASCVTACPYNYLSTDVGSCTLVC	315
Qy	334	-----	333
Db	316	PLHNQEVTAEDGTQCEKCKPCARVCYGLGMEHLREVRVTSANIQBFAGCKKIFGSLA	375
Qy	334	-----	333
Db	376	FLPESFDGDPASNTAPLQEQVFTLEITGYLYISAWPDSLPLDSVFONLQVIRGRI	435
Qy	334	-----	333
Db	436	LHNGAYSILTLOGLIGLSWLRSLRSLRSGSLAIHHNTHLCFVHTVPWDLPFRNPHQALLH	495
Qy	334	-----	333
Db	496	TANRPEDECVGEGLAGLACHQLCARGHGWPGPTQCVNCSQFLRGQECVBCRVLQGLPREYV	555
Qy	334	-----	333
Db	556	NARHCLPCHPCQPQNGSVTCFGEADQCVAACHYKDDPPFCVACRCPGVKPDLSYMPIWK	615
Qy	334	-----	333

Db	616	FPDEEGACQPCPINCTHSCVDLDDKGCAPQORASPLTISIISAVVGIILLVVLGVVFGILI	675
Qy	334	-----	333
Db	676	KRRQOKIRKYTMRRLLQETELVPLTPSGAMPNQAMRILKETELRKVKVLGSGAGFTVY	735
Qy	334	-----	333
Db	736	KGWIWPGENVKIPVAIKVIRENTSPKANKEILDEAYMAGVSPYVSRLLGICLTSTVQ	795
Qy	334	-----	333
Db	796	LVTQLMEYGLLDHVRNRCRLGSDLLNMCQIAKGSYLEVDVLRVHRLAARNVLVKS	855
Qy	334	-----	333
Db	856	PNHVKITDFGLARLLDIDETEHADGKVPKMALESILRRRTHQSDVWSYGVTVWEL	915
Qy	334	-----NPEKL---	338
Db	916	MTFGAKPYDGIPIAREIPDLLEKGERLPQPPCTIDVTYMWKMWIDSECRPRELVSE	975
Qy	339	-----	338
Db	976	FSRMARDPQRVVIQNEIDLGPASPLDSTFYRSLLDDDDMGDLVDABEYLVPQQGFFCPDP	1035
Qy	339	---GAGGMVHRHSSSTRSGGDLTLGLEPSEAPSPAPSGAGSDVFDGLGMGAA	396
Db	1036	APGAGGMVHRHSSSTRSGGDLTLGLEPSEAPSPAPSGAGSDVFDGLGMGAA	1095
Qy	397	KGQLSLPHTDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYVNPQVRRPSPREG	456
Db	1096	KGQLSLPHTDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYVNPQVRRPSPREG	1155
Qy	457	PLPAARPAAGATLRAKTLSPGKGVKVDVAFGAVENPEYLTPOGGAAPQHPHPPAFSP	516
Db	1156	PLPAARPAAGATLRAKTLSPGKGVKVDVAFGAVENPEYLTPOGGAAPQHPHPPAFSP	1215
Qy	517	AFNLVYWDODPPERGAPPSTFKGTPTAENPEYLGLDVP	555
Db	1216	AFNLVYWDODPPERGAPPSTFKGTPTAENPEYLGLDVP	1254

RESULT 2
TVRTNU
Protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 11-Jun-1999
C;Accession: A24562; A61204
R;Bargmann, C.I.; Hung, M.C.; Weinberg, R.A.
Nature 319, 226-230, 1986
A;Title: The neu oncogene encodes an epidermal growth factor receptor-related protein.
A;Reference number: A24562; UID:86118662; PMID:3945311
A;Accession: A24562
A;Molecule type: mRNA
A;Residues: 1-1260 <BAR>
A;Cross-references: EMBL:X03362; NID:G56745; PIDN:CAA27059.1; PID:G56746
R;Masul, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohen,
Carcinogenesis 12, 1975-1978, 1991
A;Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals no m
2-thiazolyl]formamide or N-methyl-N-nitrosourea.
A;Reference number: A61204; UID:92035293; PMID:1682063
A;Accession: A61204
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 637-663, 'V', 665-702 <MAS>
A;Note: authors translated the codon GCA for residue 25 as Val
C;Genetics:
A;Gene: neu
C;Superfamily: epidermal growth factor receptor; protein kinase homology
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:41:37 ; Search time 73.7051 Seconds
(without alignments)
3032.635 Million cell updates/sec

Title: US-09-821-883-4

Perfect score: 3783

Sequence: 1 MRAAPLLARAAASLSGLFLF.....CWKPVQKGA PPPRAHHHHH 697

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3783	100.0	697	US-09-821-883-4	Sequence 4, Appli
2	3473.5	91.8	690	US-09-821-883-2	Sequence 2, Appli
3	3014	79.7	564	US-09-821-883-3	Sequence 3, Appli
4	2957.5	78.2	555	US-09-821-883-1	Sequence 1, Appli
5	2585	68.3	919	US-09-854-356-6	Sequence 6, Appli
6	2478	65.5	479	US-09-821-883-5	Sequence 5, Appli
7	2422	64.0	1255	US-09-811-123-9	Sequence 9, Appli
8	2422	64.0	1255	US-09-811-115-3	Sequence 3, Appli
9	2422	64.0	1255	US-09-769-508-2	Sequence 2, Appli
10	2422	64.0	1255	US-09-984-092-4	Sequence 4, Appli
11	2422	64.0	1255	US-10-177-293-126	Sequence 136, App
12	2422	64.0	1255	US-10-207-498-6	Sequence 6, Appli
13	2422	64.0	1255	US-10-338-730-2	Sequence 2, Appli
14	2422	64.0	1255	US-10-322-892-4	Sequence 4, Appli
15	2422	64.0	1255	US-10-272-437A-28	Sequence 28, Appli

16	2422	64.0	1255	15	US-10-117-937-594	Sequence 594, App
17	2422	64.0	1255	15	US-10-435-696-36	Sequence 36, Appli
18	2422	64.0	1255	16	US-10-441-779C-4	Sequence 4, Appli
19	2422	64.0	1255	16	US-10-734-564-126	Sequence 136, App
20	2417	63.9	1255	9	US-09-854-356-1	Sequence 1, Appli
21	2417	63.9	1255	9	US-09-930-125-2	Sequence 2, Appli
22	2417	63.9	1255	10	US-09-441-411-6	Sequence 6, Appli
23	2417	63.9	1255	12	US-10-469-162-3	Sequence 3, Appli
24	2417	63.9	1255	12	US-10-253-286-553	Sequence 553, App
25	2417	63.9	1255	12	US-09-765-973-2	Sequence 2, Appli
26	2417	63.9	1255	12	US-10-418-027-3	Sequence 3, Appli
27	2417	63.9	1255	14	US-10-207-555-45	Sequence 45, Appli
28	2417	63.9	1255	14	US-10-313-644-2	Sequence 2, Appli
29	2417	63.9	1255	15	US-10-394-322A-17	Sequence 17, Appli
30	2417	63.9	1255	15	US-10-245-871-553	Sequence 553, App
31	2417	63.9	1255	15	US-10-149-138-4641	Sequence 4641, Ap
32	2417	63.9	1255	16	US-10-647-005-68	Sequence 68, Appli
33	2417	63.9	1255	16	US-10-149-138-4641	Sequence 4641, Ap
34	2410	63.7	1253	14	US-10-146-473-72	Sequence 72, Appli
35	1974.5	52.2	1256	9	US-09-854-356-14	Sequence 118, App
36	1962	51.9	1260	9	US-09-870-759-118	Sequence 118, App
37	1962	51.9	1260	10	US-09-751-708A-118	Sequence 2, Appli
38	1959.5	51.8	1256	9	US-09-854-356-2	Sequence 1, Appli
39	1614.5	42.7	645	9	US-09-921-161-1	Sequence 13, Appli
40	1614.5	42.7	645	14	US-10-368-501-13	Sequence 13, Appli
41	1614.5	42.7	645	15	US-10-608-626-13	Sequence 3, Appli
42	1614.5	42.7	653	9	US-09-854-356-3	Sequence 3, Appli
43	1614.5	42.7	685	15	US-10-412-804A-4	Sequence 4, Appli
44	1614.5	42.7	690	15	US-10-412-804A-11	Sequence 11, Appli
45	1614.5	42.7	712	9	US-09-854-356-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1

US-09-821-883-4
; Sequence 4, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; TITLE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821.883
; CURRENT FILING DATE: 2001-03-30
; PRIOR FILING DATE: 2001-03-30
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500*-rGM-CSF construct
US-09-821-883-4

Query Match 100.0%; Score 3783; DB 9; Length 697;
Best Local Similarity 100.0%; Pred. No. 2e-258;
Matches 697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MRAAPLLARAAASLSGLFLFPWLDRSVLAKELARGAASQTCTGDMKRLPASSET	60
Db	1	MRAAPLLARAAASLSGLFLFPWLDRSVLAKELARGAASQTCTGDMKRLPASSET	60
Qy	61	HLDMRLHYQCQVQVQGNLELYLPTNASLSFLQDIEVQGVYLIHNOVQVPLQRLRI	120
Db	61	HLDMRLHYQCQVQVQGNLELYLPTNASLSFLQDIEVQGVYLIHNOVQVPLQRLRI	120
Qy	121	VRGTQLPEDNYALVDNGDPLNNTPVTGASPGGLRELQRLSUTETILKGGVLIQNPOL	180

```
Db 121 VRTQLFEDYALAVLNDGDPNNPTPTVGTASPGGLRELQLRSLEILKGGVLQRNPQL 180
Qy 181 CYQDTILWKDIFHKNNQALFLIDTNRSRACHPCSPMCKGSRGWGESSEDCQSILTRTVCA 240
Db 181 CYQDTILWKDIFHKNNQALFLIDTNRSRACHPCSPMCKGSRGWGESSEDCQSILTRTVCA 240
Qy 241 GGCARCKGPLETDCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFES 300
Db 241 GGCARCKGPLETDCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFES 300
Qy 301 MPNPEGRTYTFGASCVTACPNYVNSTDVGSASIIINFEKLGAGGMVHRRHRSSTSRGGDL 360
Db 301 MPNPEGRTYTFGASCVTACPNYVNSTDVGSASIIINFEKLGAGGMVHRRHRSSTSRGGDL 360
Qy 361 TLGLEPSEEEAPRPLAPSEGAGSDVFDGLGMGAAGLQSLPHTDPSPLQRYSEDPTVP 420
Db 361 TLGLEPSEEEAPRPLAPSEGAGSDVFDGLGMGAAGLQSLPHTDPSPLQRYSEDPTVP 420
Qy 421 LPSETDGVAPLTCSPQPEYVNDVRPQPPSPREGPLPAARPAAGATLERAKTILSPGKNG 480
Db 421 LPSETDGVAPLTCSPQPEYVNDVRPQPPSPREGPLPAARPAAGATLERAKTILSPGKNG 480
Qy 481 VVKDVFAGGAVENPEYLTPOGGAAPQHPHPPAPSPAFDNLVYWDQDPPPERGAPPSTFKG 540
Db 481 VVKDVFAGGAVENPEYLTPOGGAAPQHPHPPAPSPAFDNLVYWDQDPPPERGAPPSTFKG 540
Qy 541 TPTAENPEYLGDPVAPAAAPTRSPNPTVTRPKHVDVDAIKEALSLLNDMRALENEKNEVDII 600
Db 541 TPTAENPEYLGDPVAPAAAPTRSPNPTVTRPKHVDVDAIKEALSLLNDMRALENEKNEVDII 600
Qy 601 SNEFSIORPTCVQTRKLYKQGLRGNLTKLNGALTMIAHYQTCNPPTPETDCIEVTTF 660
Db 601 SNEFSIORPTCVQTRKLYKQGLRGNLTKLNGALTMIAHYQTCNPPTPETDCIEVTTF 660
Qy 661 EDFIKNLKGFDFDIPDCWKVPQKGAAPPPAHHHHHH 697
Db 661 EDFIKNLKGFDFDIPDCWKVPQKGAAPPPAHHHHHH 697

RESULT 2
US-09-821-883-2
; Sequence 2, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; TITLE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821.883
; PRIOR FILING DATE: 2001-03-30
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500-hGM-CSF construct
US-09-821-883-2

Query Match 91.8%; Score 3473.5; DB 9; Length 690;
Best Local Similarity 91.7%; Pred. No. 1.4e-236;
Matches 641; Conservative 19; Mismatches 28; Indels 11; Gaps 2;
```

```
Qy 1 MRAAPLLIARAASLSLGLFLLFFWLDRLSVLAKELARGAASQVCTGTDMLRLPASPET 60
Db 1 MRAAPLLIARAASLSLGLFLLFFWLDRLSVLAKELARGAASQVCTGTDMLRLPASPET 60
```

```
Qy 61 HLDMLRHLQGVQVQGNLELTYLPTNASISFLQDIQEVQGYVLIHNRQVQVPLQRLRI 120
Db 61 HLDMLRHLQGVQVQGNLELTYLPTNASISFLQDIQEVQGYVLIHNRQVQVPLQRLRI 120
Qy 121 VRTQLFEDYALAVLNDGDPNNPTPTVGTASPGGLRELQLRSLEILKGGVLQRNPQL 180
Db 121 VRTQLFEDYALAVLNDGDPNNPTPTVGTASPGGLRELQLRSLEILKGGVLQRNPQL 180
Qy 181 CYQDTILWKDIFHKNNQALFLIDTNRSRACHPCSPMCKGSRGWGESSEDCQSILTRTVCA 240
Db 181 CYQDTILWKDIFHKNNQALFLIDTNRSRACHPCSPMCKGSRGWGESSEDCQSILTRTVCA 240
Qy 241 GGCARCKGPLETDCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFES 300
Db 241 GGCARCKGPLETDCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFES 300
Qy 301 MPNPEGRTYTFGASCVTACPNYVNSTDVGSASIIINFEKLGAGGMVHRRHRSSTSRGGDL 360
Db 301 MPNPEGRTYTFGASCVTACPNYVNSTDVGSASIIINFEKLGAGGMVHRRHRSSTSRGGDL 360
Qy 361 TLGLEPSEEEAPRPLAPSEGAGSDVFDGLGMGAAGLQSLPHTDPSPLQRYSEDPTVP 420
Db 361 TLGLEPSEEEAPRPLAPSEGAGSDVFDGLGMGAAGLQSLPHTDPSPLQRYSEDPTVP 420
Qy 421 LPSETDGVAPLTCSPQPEYVNDVRPQPPSPREGPLPAARPAAGATLERAKTILSPGKNG 480
Db 421 LPSETDGVAPLTCSPQPEYVNDVRPQPPSPREGPLPAARPAAGATLERAKTILSPGKNG 480
Qy 481 VVKDVFAGGAVENPEYLTPOGGAAPQHPHPPAPSPAFDNLVYWDQDPPPERGAPPSTFKG 540
Db 481 VVKDVFAGGAVENPEYLTPOGGAAPQHPHPPAPSPAFDNLVYWDQDPPPERGAPPSTFKG 540
Qy 541 TPTAENPEYLGDPVAPAAAPTRSPNPTVTRPKHVDVDAIKEALSLLNDMRALENEKNEVDII 600
Db 541 TPTAENPEYLGDPVAPAAAPTRSPNPTVTRPKHVDVDAIKEALSLLNDMRALENEKNEVDII 600
Qy 591 SNEFSIORPTCVQTRKLYKQGLRGNLTKLNGALTMIAHYQTCNPPTPETDCIEVTTF 660
Db 591 SNEFSIORPTCVQTRKLYKQGLRGNLTKLNGALTMIAHYQTCNPPTPETDCIEVTTF 660
Qy 661 EDFIKNLKGFDFDIPDCWKVPQKGAAPPPAHHHHHH 697
Db 661 EDFIKNLKGFDFDIPDCWKVPQKGAAPPPAHHHHHH 697

RESULT 3
US-09-821-883-3
; Sequence 3, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; TITLE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821.883
; PRIOR FILING DATE: 2001-03-30
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500* construct
US-09-821-883-3

Query Match 79.7%; Score 3014; DB 9; Length 564;
Best Local Similarity 100.0%; Pred. No. 3e-204;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:31:25 ; Search time 20.3037 Seconds
(without alignments)
1772.256 Million cell updates/sec

Title: US-09-821-883-4

Perfect score: 3783

Sequence: 1 MRAAPLLARASLSGLFL.....CWKPVQKAPPPPAHHHHH 697

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pap.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pap.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pap.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pap.*
5: /cgn2_6/ptodata/2/1aa/PTUS_COMB.pap.*
6: /cgn2_6/ptodata/2/1aa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2422	64.0	1255	4	US-09-527-487-2
2	2422	64.0	1255	4	US-09-811-115-3
3	2417	63.9	1255	1	US-08-467-083-68
4	2417	63.9	1255	1	US-08-414-417B-68
5	2417	63.9	1255	2	US-08-486-348A-68
6	2417	63.9	1255	2	US-08-625-101-2
7	2417	63.9	1255	2	US-08-468-545B-68
8	2417	63.9	1255	2	US-08-356-786-2
9	2417	63.9	1255	3	US-08-466-680B-68
10	2417	63.9	1255	4	US-09-354-533-68
11	2397	63.4	1255	2	US-08-484-438-8
12	2001.5	52.9	782	3	US-09-146-283-4
13	2001.5	52.9	782	3	US-08-579-823A-4
14	2001.5	52.9	782	3	US-09-344-195-4
15	1610.5	42.6	419	4	US-09-630-155-2
16	1599.5	42.3	624	3	US-08-422-108-1
17	1599.5	42.3	624	4	US-08-422-734-1
18	1177	31.1	580	1	US-08-414-417B-69
19	1177	31.1	580	2	US-08-486-348A-69
20	1177	31.1	580	2	US-08-468-545B-69
21	1177	31.1	580	3	US-08-466-680B-69
22	1177	31.1	580	4	US-09-354-533-69
23	860	22.7	166	4	US-09-648-067A-1
24	723.5	19.1	911	2	US-08-484-438-10
25	723.5	19.1	1058	2	US-08-484-438-4
26	723.5	19.1	1308	2	US-08-484-438-2
27	719.5	19.0	644	1	US-08-336-708A-9

28	719.5	19.0	1210	2	US-08-484-438-7	Sequence 7, Appli
29	719.5	19.0	1210	2	US-08-475-035-4	Sequence 4, Appli
30	701.5	18.5	478	4	US-09-570-454-2	Sequence 2, Appli
31	701.5	18.5	478	4	US-09-867-521-2	Sequence 2, Appli
32	697	18.4	1342	1	US-07-978-895-4	Sequence 4, Appli
33	697	18.4	1342	2	US-08-484-438-9	Sequence 9, Appli
34	697	18.4	1342	2	US-08-473-119-4	Sequence 4, Appli
35	697	18.4	1342	2	US-08-475-352-4	Sequence 4, Appli
36	697	18.4	1342	4	US-09-170-699-4	Sequence 4, Appli
37	697	18.4	1343	6	5183884-4	Patent No. 5183884
38	529.5	14.0	515	2	US-09-146-283-2	Sequence 2, Appli
39	529.5	14.0	515	3	US-08-579-823A-2	Sequence 2, Appli
40	529.5	14.0	515	3	US-09-344-195-2	Sequence 2, Appli
41	493	13.0	97	1	US-08-421-356-3	Sequence 3, Appli
42	493	13.0	97	1	US-09-046-783-3	Sequence 3, Appli
43	472.5	12.5	219	2	US-08-902-516-2	Sequence 2, Appli
44	472.5	12.5	219	4	US-09-847-185-2	Sequence 2, Appli
45	469	12.4	141	1	US-08-259-696B-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1

US-09-527-487-2
; Sequence 2, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126881309200
; CURRENT APPLICATION NUMBER: US/09/527.487
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-527-487-2

Query Match	64.0%	Score	2422	DB	4	Length	1255
Best Local Similarity	41.3%	Pred. No.	2.2e-187				
Matches	512	Conservative	4	Mismatches	5	Indels	718
							Gap 2
Qy	35	LARGAASQVCTGDMKRLPASPEHLDMLRHLYQGCQVQGNLELTYLPTNASLSFLQ	94				
Db	16	LPPGAASQVCTGDMKRLPASPEHLDMLRHLYQGCQVQGNLELTYLPTNASLSFLQ	75				
Qy	95	DIQEVQGYVLIHNVQRQVPLQRLRIVRGTQLPEDNYALAVLDNGDPLNNTTPVTGASPG	154				
Db	76	DIQEVQGYVLIHNVQRQVPLQRLRIVRGTQLPEDNYALAVLDNGDPLNNTTPVTGASPG	135				
Qy	155	GLRELQRLSLEILKGGVLTORNQOLCVQDTILWKDIFHKONQALATLIDNRSRACHPC	214				
Db	136	GLRELQRLSLEILKGGVLIQORNQOLCVQDTILWKDIFHKONQALATLIDNRSRACHPC	195				
Qy	215	SPMKGRGRCWGESSEDCQSLTRTCAGCARCKGKPLPTDCCHCCAGCTGPKHSDCLAC	274				
Db	196	SPMKGRGRCWGESSEDCQSLTRTCAGCARCKGKPLPTDCCHCCAGCTGPKHSDCLAC	255				
Qy	275	LHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPYNYLSTDVGSASII-	333				
Db	256	LHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPYNYLSTDVGSCTLVC	315				
Qy	334	-----	333				
Db	316	PLHNQVETADGTQRCCKSPCARVCYGLGMEHLREVRVTSANIOFAGCKKIFGSLA	375				
Qy	334	-----	333				
Db	376	FLPESFGDPPASNTAPLQPEQLQVFETLEETGYLYISAWPDSLPDLSVFNQLQVIRGRI	435				

```
Qy 334 ----- 333
Db 436 LHNAYSLTLOGLISWLGSLRLSRLGSLALIHNNTHLCFVHTVPWDQFRNPQHALLH 495
Qy 334 ----- 333
Db 496 TANRPEDECYVGEGLACHQLCARGHCWGPGPTQCVCNCSQFLRGQCEBECRVLOGLPREYV 555
Qy 334 ----- 333
Db 556 NARHCLPCHPECPQNGSVTCFGEADQCVACAHYKDPFCVACPSGVKPDLSYMPIWK 615
Qy 334 ----- 333
Db 616 FPDEGACQPCPNCTHSCVDLDDKGCPCAPORASPLTSIVSAVVGILLVVVLGVVFGILI 675
Qy 334 ----- 333
Db 676 KRRQKIRKYTMRRLLQETELVEPLTPSGAMPNQAMRILKETELRKVKVLGSGAFCTVY 735
Qy 334 ----- 333
Db 736 KGIWPDGENVKIPVAIKVLRNTPSKANKEILDEAYMAGVGSYVSRLLGICLTSTVQ 795
Qy 334 ----- 333
Db 796 LVTQMPYGCLLDHRVNRGRGLSQDILLNMCQIAKMSYLEDLVLRHDLAARNVLKS 855
Qy 334 ----- 333
Db 856 PNHVKITDFGLARLLDIDETEHADGGKVPKIMWMALESILRRRPTHQSDVMSYGVTVWEL 915
Qy 334 ----- 338
Db 916 MTFCAKPYDGIPAREIPDLLEKGERLPQPICTIDVYIMVIMWCMIDSECRPRELVS 975
Qy 339 ----- 338
Db 976 FSRWAPQRFVIVQNEDELGPASPLDSTFYRSLLDDMDGLVDAEYLVPOQGFCDP 1035
Qy 339 --GAGGVHRRSSSTRSGGDLTLGLEPSEEBAPRSPAPSEAGSDVPDGLMGAA 396
Db 1036 APGAGGVHRRSSSTRSGGDLTLGLEPSEEBAPRSPAPSEAGSDVPDGLMGAA 1095
Qy 397 KGLQSLTHDPSPLQRYSEDTPVLPSETDGYVAPLTCSPQPEYVQPDVVRPQPPSPREG 456
Db 1096 KGLQSLTHDPSPLQRYSEDTPVLPSETDGYVAPLTCSPQPEYVQPDVVRPQPPSPREG 1155
Qy 457 PLPAARPAATLERAKTILSPKNGVWVDVAFGAVENPEYLTPOQGAAPOPHPPPAFSP 516
Db 1156 PLPAARPAATLERAKTILSPKNGVWVDVAFGAVENPEYLTPOQGAAPOPHPPPAFSP 1215
Qy 517 AFDNLYWQDQPPRGPAPPSTFKGTPTAENPEYLGLDVP 555
Db 1216 AFDNLYWQDQPPRGPAPPSTFKGTPTAENPEYLGLDVP 1254
```

RESULT 2

```
US-09-811-115-3
; Sequence 3, Application US/09811115
; Patent No. 6632979
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A
; CURRENT APPLICATION NUMBER: US/09/811.115
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 3
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-115-3

Query Match 64.0%; Score 2422; DB 4; Length 1255;
Best Local Similarity 41.3%; Pred. No. 2.2e-187;
Matches 512; Conservative 4; Mismatches 5; Indels 718; Gaps 2;

Qy 35 LARGAASQVCTGTGDMKRLPASPETHLDMRLHYQCGVQVQGNLELYLPTNASLSFLQ 94
Db 16 LPFGAASQVCTGTGDMKRLPASPETHLDMRLHYQCGVQVQGNLELYLPTNASLSFLQ 75
Qy 95 DIQEVQGYVLIANNQROVPLQRLRIVRGTQLEPDNYALAVLDNGDPLNNTTPTVTGASPG 154
Db 76 DIQEVQGYVLIANNQROVPLQRLRIVRGTQLEPDNYALAVLDNGDPLNNTTPTVTGASPG 135
Qy 155 GLRELOLRSLTEILKGGVLIQRNPQLCYQDTILWKDIFHKNQNLALTLIDTNRSRACHPC 214
Db 136 GLRELOLRSLTEILKGGVLIQRNPQLCYQDTILWKDIFHKNQNLALTLIDTNRSRACHPC 195
Qy 215 SPMCKSRGCGESSEDCQSLTRTVACGACARCKGPLPTDCHEQCAAGCTGPKHSDCLAC 274
Db 196 SPMCKSRGCGESSEDCQSLTRTVACGACARCKGPLPTDCHEQCAAGCTGPKHSDCLAC 255
Qy 275 LHFNHSGICELHCPALVYNTDTFESMPNPEGRYTFGASCVTACPNYLSLTDVGSASII - 333
Db 256 LHFNHSGICELHCPALVYNTDTFESMPNPEGRYTFGASCVTACPNYLSLTDVGSCTLVC 315
Qy 334 ----- 333
Db 316 PLHNQEVTAEDGTQRCCKSPCARVCYGLGMEHLREVRAVTSANIOFAGCKKIFGSLA 375
Qy 334 ----- 333
Db 376 FLPEFDGDPASNTAPLQPEQLQVFETLEBITGYLISAWPDSILPDLASFQNLQVIRGRI 435
Qy 334 ----- 333
Db 436 LHNAYSLTLOGLISWLGSLRLSRLGSLALIHNNTHLCFVHTVPWDQFRNPQHALLH 495
Qy 334 ----- 333
Db 496 TANRPEDECYVGEGLACHQLCARGHCWGPGPTQCVCNCSQFLRGQCEBECRVLOGLPREYV 555
Qy 334 ----- 333
Db 556 NARHCLPCHPECPQNGSVTCFGEADQCVACAHYKDPFCVACPSGVKPDLSYMPIWK 615
Qy 334 ----- 333
Db 616 FPDEGACQPCPNCTHSCVDLDDKGCPCAPORASPLTSIVSAVVGILLVVVLGVVFGILI 675
Qy 334 ----- 333
Db 676 KRRQKIRKYTMRRLLQETELVEPLTPSGAMPNQAMRILKETELRKVKVLGSGAFCTVY 735
Qy 334 ----- 333
Db 736 KGIWPDGENVKIPVAIKVLRNTPSKANKEILDEAYMAGVGSYVSRLLGICLTSTVQ 795
Qy 334 ----- 333
Db 796 LVTQMPYGCLLDHRVNRGRGLSQDILLNMCQIAKMSYLEDLVLRHDLAARNVLKS 855
Qy 334 ----- 333
Db 856 PNHVKITDFGLARLLDIDETEHADGGKVPKIMWMALESILRRRPTHQSDVMSYGVTVWEL 915
Qy 334 ----- 338
Db 916 MTFCAKPYDGIPAREIPDLLEKGERLPQPICTIDVYIMVIMWCMIDSECRPRELVS 975
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:20:15 ; Search time 76.7646 Seconds
(without alignments)
2565.449 Million cell updates/sec

Title: US-09-821-883-4

Perfect score: 3783

Sequence: 1 MRAAPLLARAASLSGLF.....CWKPVQKAGPPPPAHHHHH 697

Scoring table: BLOSUM62 , Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003s.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3783	100.0	697	4	AAE13111 Human HER
2	3473.5	91.8	690	4	AAE13109 Human HER
3	3014	79.7	564	4	AAE13110 Human HER
4	2957.5	78.2	555	4	AAE13108 Human HER
5	2585	68.3	919	3	AB21203 Human HER
6	2585	68.3	919	5	AA51148 Her-2/neu
7	2478	65.5	479	4	AAE13112 Human HER
8	2422	64.0	1255	3	AA92620 Human her
9	2422	64.0	1255	4	AAE13111 Human HER
10	2422	64.0	1255	4	AAE12130 Human tyr
11	2422	64.0	1255	5	AAE26349 Human HER
12	2422	64.0	1255	5	AAE26366 Human HER
13	2422	64.0	1255	5	AAU74545 Human HER
14	2422	64.0	1255	6	ABR4747 Breast ca
15	2422	64.0	1255	6	ABP74708 Human HER
16	2422	64.0	1255	6	AAE38390 Human C-e
17	2422	64.0	1255	6	ADA38143 Human erb
18	2422	64.0	1255	7	ADA37255 Human Erb
19	2422	64.0	1255	7	ADB67621 Human epi
20	2420	64.0	1433	2	AAE39568 Sequence
21	2417	63.9	1255	2	AAW01111 HER-2/neu
22	2417	63.9	1255	2	AAW92406 Human HER
23	2417	63.9	1255	3	AA984780 Amino aci
24	2417	63.9	1255	3	AAE21198 Human HER
25	2417	63.9	1255	4	AAE88267 HER2/neu

ALIGNMENTS

RESULT 1

AAE13111	AAE13111 standard; protein; 697 AA.
XX	AC AAE13111;
XX	DT 28-JAN-2002 (first entry)
XX	DE Human HER500-rGM-CSF fusion construct comprising OVA-derived peptide.
XX	KW Immunostimulatory fusion protein; IFP; antigen component; therapy;
KW	Immunostimulatory component; T-cell mediated immune response; DC;
KW	dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW	PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW	membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
KW	HER-2 protein; granulocyte-macrophage colony stimulating factor;
KW	ovalbumin-derived octapeptide; OVA; rat; HER500-rGM-CSF fusion protein.
XX	OS Homo sapiens.
OS	Rattus norvegicus.
OS	Unidentified.
OS	Chimeric.
XX	PN WO200174855-A2.
XX	PD 11-OCT-2001.
XX	PF 30-MAR-2001; 2001WO-US010515.
XX	PR 30-MAR-2000; 2000US-0193504P.
XX	PA (DEND-) DENDREON CORP.
XX	PI Laus R, Vidovic D, Graddis T;
XX	DR WPI; 2001-662965/76.
XX	DR N-PSDB; AAD21567.
PT	An immunostimulatory fusion protein comprising the intracellular domain
PT	of HER-2 and an antigen elicits an immune response to the antigen and is
PT	useful for the treatment of associated cancer associated.
PS	Claim 7; Page 27; 59pp; English.
XX	The invention relates to immunostimulatory fusion proteins (IFP) and
CC	nucleic acid molecules encoding such proteins. The IFPs comprise a
CC	polypeptide antigen component and an immunostimulatory component derived
CC	from the intracellular domain of HER-2 protein which is effective to

CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
CC immune response to the antigen. IFF or superactivated dendritic cells are
CC used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
CC associated with a particularly antigen. The present sequence is HER500
CC rGM-CSF fusion protein construct which comprises human PAP signal
CC sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal
CC sequence, mature HER-2 membrane distal extracellular domain, an Ala
CC linker, an ovalbumin (OVA)-derived immunodominant octapeptide, HER-2
CC membrane distal intracellular domain, an Ala Ala linker, a mature rat
CC granulocyte-macrophage colony stimulating factor (GM-CSF) sequence and a
CC C-terminal tag
XX
SQ Sequence 697 AA;

Query Match 100.0%; Score 3783; DB 4; Length 697;
Best Local Similarity 100.0%; Pred. No. 6.2e-243;
Matches 697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAAPLLAARASLSGLFLFLFWLDRSVLAKELARGAASSTVCTGDMKRLRPASPET 60
Db 1 MRAAPLLAARASLSGLFLFLFWLDRSVLAKELARGAASSTVCTGDMKRLRPASPET 60

Qy 61 HLDMLRHLQYQCVVQGNLELYLPTNASLSFLQDIQEVQGVLIHNVQVPLQRLRI 120
Db 61 HLDMLRHLQYQCVVQGNLELYLPTNASLSFLQDIQEVQGVLIHNVQVPLQRLRI 120

Qy 121 VRGTQLFEDNYALAVLNDGDPPLNNTPTVTGASPGGLRELQRLRSITLILKGVLIQRNPQL 180
Db 121 VRGTQLFEDNYALAVLNDGDPPLNNTPTVTGASPGGLRELQRLRSITLILKGVLIQRNPQL 180

Qy 181 CYQDTILWKDIFHKNNQALTLIDNRSRACHPCSPMCKGRCWGESSEDCQSLTRTVCA 240
Db 181 CYQDTILWKDIFHKNNQALTLIDNRSRACHPCSPMCKGRCWGESSEDCQSLTRTVCA 240

Qy 241 GGCARCKGPLETDCHECAAGCTGPKHSDCLACLHFNHSGICELHCPALVYNTDTPES 300
Db 241 GGCARCKGPLETDCHECAAGCTGPKHSDCLACLHFNHSGICELHCPALVYNTDTPES 300

Qy 301 MPNPEGRYTFGASCVTACPNYLSLTDVGSASIIINFEKLGAGGMVHRRSSTRSGGDL 360
Db 301 MPNPEGRYTFGASCVTACPNYLSLTDVGSASIIINFEKLGAGGMVHRRSSTRSGGDL 360

Qy 361 TLGLEPSEEAAPRSLAPSEAGSDVFDGLGMGAAGLQSLPHTDPSPLQRYSEDPTVP 420
Db 361 TLGLEPSEEAAPRSLAPSEAGSDVFDGLGMGAAGLQSLPHTDPSPLQRYSEDPTVP 420

Qy 421 LPSETDGYVAPLTCSPQPEYVNPQDVRPQPPSPREGPLPAARPAATLAKTLSPGKNG 480
Db 421 LPSETDGYVAPLTCSPQPEYVNPQDVRPQPPSPREGPLPAARPAATLAKTLSPGKNG 480

Qy 481 VVKDVFAGGAVENPEYLTPOGGAAPQHPHPPAPSPAFDNLVYWDQDPPERGAPPSTFKG 540
Db 481 VVKDVFAGGAVENPEYLTPOGGAAPQHPHPPAPSPAFDNLVYWDQDPPERGAPPSTFKG 540

Qy 541 TPTAENPEYLGIDVPAAPATPSNPVTRPKHVDVDAIKEALSILNDMRALENEKNEVDII 600
Db 541 TPTAENPEYLGIDVPAAPATPSNPVTRPKHVDVDAIKEALSILNDMRALENEKNEVDII 600

Qy 601 SNEFSIORPTCVQTRKLYKQGLRGNLTAKNGALTMIAHYQTNCPPTDCEIEVTTF 660
Db 601 SNEFSIORPTCVQTRKLYKQGLRGNLTAKNGALTMIAHYQTNCPPTDCEIEVTTF 660

Qy 661 EDFIKNLKGFIDIPDFCWKPVQKGAAPPPPAHHHHH 697
Db 661 EDFIKNLKGFIDIPDFCWKPVQKGAAPPPPAHHHHH 697

RESULT 2
AAE13109 standard; protein; 690 AA.
XX
AC AAE13109;
XX

DT 28-JAN-2002 (first entry)
XX Human HER500-hGM-CSF fusion protein construct.
DE Immunostimulatory fusion protein; IFF; antigen component; therapy;
XX immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
KW HER-2 protein; granulocyte-macrophage colony stimulating factor;
KW HER500-hGM-CSF fusion protein.
XX
OS Homo sapiens.
OS Synthetic.
OS Chimeric.
XX WO200174855-A2.
PN 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US010515.
XX 30-MAR-2000; 2000US-0193504P.
XX (DEND-) DENDREON CORP.
XX Laus R, Vidovic D, Graddis T;
XX WPT; 2001-662965/76.
DR N-PSDB; AAD21565.
XX An immunostimulatory fusion protein comprising the intracellular domain
XX of HER-2 and an antigen elicits an immune response to the antigen and is
XX useful for the treatment of associated cancer associated.
XX Claim 7; Page 26; 59pp; English.
XX The invention relates to immunostimulatory fusion proteins (IFF) and
XX nucleic acid molecules encoding such proteins. The IFFs comprise a
XX polypeptide antigen component and an immunostimulatory component derived
XX from the intracellular domain of HER-2 protein which is effective to
XX elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
XX immune response to the antigen. IFF or superactivated dendritic cells are
XX used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
XX associated with a particularly antigen. The present sequence is HER500
XX hGM-CSF fusion protein construct which comprises human PAP signal
XX sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal
XX sequence, mature HER-2 membrane distal extracellular and intracellular
XX domains, an Ala Ala linker, a mature human granulocyte- macrophage colony
XX stimulating factor (GM-CSF) sequence and a C-terminal tag
XX
SQ Sequence 690 AA;

Query Match 91.8%; Score 3473.5; DB 4; Length 690;
Best Local Similarity 91.7%; Pred. No. 2.3e-222;
Matches 641; Conservative 19; Mismatches 28; Indels 11; Gaps 2;

Qy 1 MRAAPLLAARASLSGLFLFLFWLDRSVLAKELARGAASSTVCTGDMKRLRPASPET 60
Db 1 MRAAPLLAARASLSGLFLFLFWLDRSVLAKELARGAASSTVCTGDMKRLRPASPET 60

Qy 61 HLDMLRHLQYQCVVQGNLELYLPTNASLSFLQDIQEVQGVLIHNVQVPLQRLRI 120
Db 61 HLDMLRHLQYQCVVQGNLELYLPTNASLSFLQDIQEVQGVLIHNVQVPLQRLRI 120

Qy 121 VRGTQLFEDNYALAVLNDGDPPLNNTPTVTGASPGGLRELQRLRSITLILKGVLIQRNPQL 180
Db 121 VRGTQLFEDNYALAVLNDGDPPLNNTPTVTGASPGGLRELQRLRSITLILKGVLIQRNPQL 180

Qy 181 CYQDTILWKDIFHKNNQALTLIDNRSRACHPCSPMCKGRCWGESSEDCQSLTRTVCA 240
Db 181 CYQDTILWKDIFHKNNQALTLIDNRSRACHPCSPMCKGRCWGESSEDCQSLTRTVCA 240

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:57:41 ; Search time 3803.06 Seconds
(without alignments)
4428.608 Million cell updates/sec

Title: US-09-821-883-3
Perfect score: 3062
Sequence: 1 MRAAPLLARASLSGLFLP.....ENPEYLGVDVPAAHHHHH 564

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODEL=frame+ p2n.model -US9821883/runat 09092004 105127 6643/app.query.fasta_1.3100
-Q=/cgn2_1/USPTC spool/US9821883 -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-DB=EST -QFMT=fastap -SUFFIX=1 -TRANS=human40.coi -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -ALIGN=15 -MODE=LOCAL
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US9821883 @CGN 1 1 11571 @runat 09092004 105127 6643 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hcc.*
9: gb_est1.*
10: gb_est2.*
11: gb_hcc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gsa_hum.*
18: em_gsa_inv.*
19: em_gsa_pln.*
20: em_gsa_vrt.*
21: em_gsa_fun.*
22: em_gsa_mam.*
23: em_gsa_mus.*
24: em_gsa_pro.*
25: em_gsa_rod.*
26: em_gsa_phg.*
27: em_gsa_vrl.*
28: gb_gsa1.*

29: gb_gsa2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1977.5	64.6	4323	11 AK031099	AK031099 Mus muscu
2	1367	64.2	4463	11 AK083669	AK083669 Mus muscu
3	1291.5	42.2	1016	12 BM562913	BM562913 AGENCOURT
4	1169	38.2	1129	12 BM802792	BM802792 AGENCOURT
5	1157	37.8	871	14 CA488274	CA488274 AGENCOURT
6	1157	37.8	885	14 CA455074	CA455074 AGENCOURT
7	1150	37.6	4715	11 AF318349	AF318349 Homo sapi
8	1131.5	37.0	893	14 CA455141	CA455141 AGENCOURT
9	1084	35.4	1004	14 CD515356	CD515356 AGENCOURT
10	1081	35.3	894	14 CA454131	CA454131 AGENCOURT
11	1080	35.3	653	12 BM721340	BM721340 UI-E-E01-
12	1073	35.0	588	13 BX478931	BX478931 DKFZp686A
13	1072	35.0	902	14 CA488868	CA488868 AGENCOURT
14	1046	34.2	569	9 AL701765	AL701765 DKFZp686P
15	1033	33.4	583	13 BX479114	BX479114 DKFZp686P
16	1022.5	33.4	795	14 CA328613	CA328613 UI-M-FY0-
17	1021	33.3	3110	11 AK031542	AK031542 Mus muscu
18	1010	33.0	808	14 CA489534	CA489534 AGENCOURT
19	991	32.4	616	14 CD721801	CD721801 o101f07.y
20	973	31.8	1988	11 BC023725	BC023725 Mus muscu
21	972	31.7	943	13 BQ958632	BQ958632 AGENCOURT
22	931	30.4	813	12 BI557797	BI557797 603236977
23	926	30.2	855	13 BU594980	BU594980 AGENCOURT
24	921	30.1	614	10 AK370693	AK370693 QV1-BT026
25	910	29.7	717	14 CB598701	CB598701 AGENCOURT
26	909.5	29.7	852	14 CD516283	CD516283 AGENCOURT
27	908	29.7	791	12 BI154872	BI154872 602902857
28	901	29.4	491	12 BM790293	BM790293 K-EST0070
29	886	28.9	906	14 CA454570	CA454570 AGENCOURT
30	883	28.8	998	12 BI649877	BI649877 603296516
31	875	28.6	685	9 AU123871	AU123871 AGENCOURT
32	873.5	28.5	932	14 CA487981	CA487981 AGENCOURT
33	868	28.3	609	10 AW701942	AW701942 uq93b02.y
34	862	28.2	621	12 BG283493	BG283493 602407782
35	855	27.9	786	12 BI155788	BI155788 602904360
36	846	27.6	887	13 BQ717097	BQ717097 AGENCOURT
37	835.5	27.3	649	9 AA496412	AA496412 zv37802.r
38	817	26.7	533	14 CB437620	CB437620 685491 MA
39	817	26.7	552	12 BG277542	BG277542 ux45d07.y
40	817	26.7	670	9 AI906012	AI906012 RC-BT105-
41	817	26.7	876	13 BQ769889	BQ769889 UI-M-F10-
42	808	26.4	876	13 BQ769889	BQ769889 UI-M-F10-
43	805	26.3	514	10 BF998814	BF998814 QVO-GN014
44	800	26.1	887	14 CA980253	CA980253 AGENCOURT
45	797.5	26.0	678	12 BI555157	BI555157 603236396

ALIGNMENTS

RESULT 1
AK031099
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AK031099 4323 bp mRNA linear HTC 18-SEP-2003
Mus musculus 13 days embryo forelimb cDNA, RIKEN full-length
enriched library, Clone:5930404N10 product:v-erb-b2 erythroblastic
leukemia viral oncogene homolog 2, neuro/glioblastoma derived
oncogene homolog (avian), full insert sequence.
AK031099
AK031099.1 GI:26082143
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
MEDLINE
PUBMED
10349636
2 Carninci, P., Shibata, Y., Hayateu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE
PUBMED
11042159
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
PUBMED
11076861
4

The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 4323)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayateu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, C., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

Location/Qualifiers
1. .4323
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

FEATURES

source

/db_xref="PANTOM_DB:5930404N10"
/db_xref="MGI:2394277"
/db_xref="taxon:10090"
/clone="5930404N10"
/tissue_type="forelimb"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
48. .3818
misc_feature
/note="putative
v-erb-b2 erythroblastic leukemia viral oncogene homolog 2,
neuro/glioblastoma derived oncogene homolog (avian)
[MGI:MGI:95410, GB|U71126, evidence: BLASTN, 99%,
match=449]"

ORIGIN

Alignment Scores:
Pred. No.: 2,43e-114 Length: 4323
Score: 1977.50 Matches: 439
Percent Similarity: 36.72% Conservative: 31
Best Local Similarity: 34.30% Mismatches: 79
Query Match: 64.58% Indels: 731
DB: 11 Gaps: 4
US-09-821-883-3 (1-564) x AK031099 (1-4323)
Qy 3 AlaAlaProLeuLeuAlaAlaAlaSerLeuSerLeuGlyPheLeuLeuLeu 22
Db 33 GCCGAGTAAATCATCATGAGCTGGCGCTGGTGGGGTTCCTCTCGCCCTC 92
Qy 23 PhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAlaSerThr 42
Db 93 -----CTGTCCTCCGAGCGCGGTACC 116
Qy 43 GlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThrHisLeu 62
Db 117 CAAGTGTGTACCGGTACCGACATGAAGTGGATCTCCCTGCGAGTCTCGAGACCCACCTG 176
Qy 63 AspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnGlyValGlnGlyLeuThr 82
Db 177 GACATGCTTCCACCTCTACACGGCTGTGAGGGTGTGAGGGCAATTGGAGCTTACC 236
Qy 83 TyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspLleGlnGlnValGlnGlyTyr 102
Db 237 TACCTGCGCGCAATGCCAGCTCTCATTCCTGCGAGGACATCCAGGAAGTCCAGGGATAC 296
Qy 103 ValLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgGlyLeuValArg 122
Db 297 ATGCTCATCGCTCACACCGAGTGAACACACGCTCCCTGCGAGGTTGCGCATCTGTGAGA 356
Qy 123 GlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAspProLeu 142
Db 357 GGGACTCAGCTCTTTGAGGACAAGTATGCTGCTGCTGTAGACAACCGAGACCTTTTG 416
Qy 143 AsnAsn---ThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGlnLeu 161
Db 417 GACACGCTCACACCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 476
Qy 162 ArgSerLeuThrGluLeuLeuGlyValLeuLeuGlnArgAsnProGlnLeuCys 181
Db 477 CGAAGTCTCACAGAGATCTTGAAGGAGAGAGTGTGATCCGTGGGAAACCTCAGCTCTGC 536
Qy 182 TyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeuAlaLeuThr 201
Db 537 TACACGAGCATGGTTTGTGAAGAGATGCTCCGTAGAGATATACACAGTGGCTCTCTGTC 596
Qy 202 LeuLeuAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGlySer 221
Db 597 GACATGGACACCAATCGTTCCCGGCGCTGTCCACCTTGTGCGCCCAACCTGCAAGACAAT 656
Qy 222 ArgCysTrpGlyGluSerGluAspCysGlnSerLeuThrArgThrValCysAlaGly 241
Db 657 CACTGTTGGGTGAGAGTCTCTGAAGAGATGTCTGAGATCTTGACTGGCACCATCTGTACTAGT 716

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 15:10:16 ; Search time 681.931 Seconds
(without alignments)
4161.385 Million cell updates/sec

Title: US-09-821-883-3
Perfect score: 3062
Sequence: 1 MRAAPLLARAASLSGLFLF.....ENPEYLGVDVPAAAHHHHH 564

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3304383 seqs, 2515761380 residues

Total number of hits satisfying chosen parameters: 6608766

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DB=frame+ p2n model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09821883/runat_09092004_105129_6736/app.query.fasta_1.3100
-DB=Published Applications NA -Qfmt=fastap -SUFFIX=nrnp -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USERS=US09821883 @CNC 1 1 1949 @runat_09092004_105129_6736
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
------------	-------	-------	--------	----	-------------

RESULT 1
US-09-821-883-8
; Sequence 8, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1692
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500* construct

ALIGNMENTS

1	3062	100.0	1692	9	US-09-821-883-8	Sequence 8, Appl
2	3014	98.4	2091	9	US-09-821-883-9	Sequence 9, Appl
3	3005.5	96.2	1665	9	US-09-821-883-6	Sequence 6, Appl
4	2957.5	96.6	2070	9	US-09-821-883-7	Sequence 7, Appl
5	2430.5	79.4	4543	9	US-09-769-508-1	Sequence 1, Appl
6	2427.5	79.3	4530	9	US-09-877-177-11	Sequence 11, Appl
7	2427.5	79.3	4530	13	US-10-392-113-45	Sequence 45, Appl
8	2427.5	79.3	4530	13	US-10-426-836-11	Sequence 11, Appl
9	2427.5	79.3	4530	15	US-10-177-293-125	Sequence 125, App
10	2427.5	79.3	4530	15	US-10-007-926A-119	Sequence 119, App
11	2427.5	79.3	4530	15	US-10-338-730-1	Sequence 1, Appl
12	2427.5	79.3	4530	15	US-10-101-510-124	Sequence 124, App
13	2427.5	79.3	4530	16	US-10-116-275-131	Sequence 131, App
14	2427.5	79.3	4530	16	US-10-272-437A-27	Sequence 27, Appl
15	2427.5	79.3	4530	16	US-10-117-937-595	Sequence 595, App
16	2427.5	79.3	4530	16	US-10-159-563-208	Sequence 208, App
17	2427.5	79.3	4530	16	US-10-435-696-10	Sequence 10, Appl
18	2427.5	79.3	4530	17	US-10-734-564-59	Sequence 59, Appl
19	2427.5	79.3	4642	15	US-10-198-846-10896	Sequence 10896, A
20	2425	79.2	9274	9	US-09-811-123-7	Sequence 7, Appl
21	2425	79.2	9274	9	US-09-811-115-1	Sequence 1, Appl
22	2422.5	79.1	4473	10	US-09-441-411-5	Sequence 5, Appl
23	2422.5	79.1	4473	15	US-10-146-473-32	Sequence 32, Appl
24	2422.5	79.1	4473	15	US-10-207-655-44	Sequence 44, Appl
25	2422.5	79.1	4473	15	US-10-101-510-81	Sequence 81, Appl
26	2422	79.1	3765	15	US-10-207-498-5	Sequence 5, Appl
27	2422	79.1	3768	9	US-09-811-123-8	Sequence 8, Appl
28	2422	79.1	3768	9	US-09-811-115-2	Sequence 2, Appl
29	2422	79.1	3768	13	US-09-984-092-3	Sequence 3, Appl
30	2422	79.1	3768	13	US-10-280-576-3	Sequence 3, Appl
31	2422	79.1	3768	17	US-10-441-779C-3	Sequence 3, Appl
32	2417	78.9	3768	9	US-09-854-356-9	Sequence 9, Appl
33	2417	78.9	3768	9	US-09-930-125-1	Sequence 1, Appl
34	2417	78.9	3768	13	US-09-765-373-1	Sequence 1, Appl
35	2417	78.9	3768	15	US-10-313-644-1	Sequence 1, Appl
36	2402.5	78.5	4606	10	US-09-971-392-70	Sequence 70, Appl
37	1974.5	64.5	3771	9	US-09-854-356-11	Sequence 11, Appl
38	1968	64.3	3955	9	US-09-870-759-117	Sequence 117, App
39	1968	64.3	3955	9	US-09-854-356-10	Sequence 10, Appl
40	1968	64.3	3955	10	US-09-751-708A-117	Sequence 117, App
41	1838.5	60.0	1437	9	US-09-821-883-10	Sequence 10, Appl
42	1644	53.7	2132	16	US-10-412-804A-3	Sequence 3, Appl
43	1641.5	53.6	2149	16	US-10-412-804A-9	Sequence 9, Appl
44	1638.5	53.5	2164	16	US-10-412-804A-5	Sequence 5, Appl
45	1587	51.8	867	9	US-09-821-883-24	Sequence 24, Appl

US-09-821-883-8

Alignment Scores:

Pred. No.: 1,22e-282 Length: 1692
Score: 3062.00 Matches: 564
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 9

US-09-821-883-3 (1-564) x US-09-821-883-8 (1-1692)

Qy 1 MetArgAlaLaProLeuLeuLeuAlaArgAlaLaSerLeuSerLeuGlyPheLeuPhe 20
Db 1 ATGAGAGCTGACCCCTCTCTGCGCAGGAGGAGCAAGCCTTAGCCTTGGCTTCTTGT 60
Qy 21 LeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaLa 40
Db 61 CTGCTTTTCTGGCTAGACCGAAGTGTACTAGCCAAAGAGTGTGGCGCGCGGCGCG 120
Qy 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
Db 121 TCGACCCCAAGTGTGACCGGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAGACC 180
Qy 61 HisLeuAspMetLeuArgHisLeuTrpGlnGlyCysGlnValValGlnGlyAsnLeuGlu 80
Db 181 CACCTGGACATGCTCCGCACTCTACCCAGGGCTGCCAGGTGGTGCGAGGAAACCTGGAA 240
Qy 81 LeuThrTrpLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGln 100
Db 241 CTCACTACCTGCCCAACATGCCAGCTGCTCTCTCCCTGCGAGATATCCAGGAGGTGCAG 300
Qy 101 GlyTrpValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgTle 120
Db 301 GGCTACGTGCTCATGCTGCTCAACCAAGTAGGAGGAGTCCCACTGCAGAGGCTCGGAT 360
Qy 121 ValArgGlyThrGlnLeuPheGluAspAsnTrpAlaLeuAlaValLeuAspAsnGlyAsp 140
Db 361 GTGCGAGGCAACCCAGCTCTTTGAGGACAACTATGCCCTGGCCGTCTAGACAATGGAGAC 420
Qy 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln 160
Db 421 CCGCTGAAACAATACCAACCTGTACAGGGGCTCCCAAGGAGGCTCGGGAGGTGCAG 480
Qy 161 LeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeu 180
Db 481 CTTCGAGCCCTCACAGAGATCTTGAAGAGAGGGGTCTTGATCCAGCGGAAACCCAGCTC 540
Qy 181 CysTrpGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200
Db 541 TGCTACCAAGGACACGATTTTGTGGAGGAGCATCTTCCACAGNACAACACAGCTGGCTCTC 600
Qy 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220
Db 601 ACATGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGTAAGGSC 660
Qy 221 SerArgCysTrpGlyLysSerSerGluAspCysGlnSerLeuThrArgThrValCysAla 240
Db 661 TCCCGCTCTCGGGAGAGAGTCTGAGGATGTGAGAGCCTGACGCGCACTGTCTGTGCC 720
Qy 241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCysAla 260
Db 721 GGTGGCTGTGCGCTGCAAGGGGCACTGCCCACTGACTGTGCTGCAATGAGCAGTGTCT 780
Qy 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
Db 781 GCCGCTGACCGGGCCCCAAGCACTCTGACTGCTGGCTGCTCCACTTCAACCCACAGT 840
Qy 281 GlyIleCysGluLeuHisCysProAlaLeuValThrTrpAsnThrAspThrPheGluSer 300
Db 841 GGCATCTGTGAGTGCATGCCACGCTGCGTGTGTCACCTACACACAGACAGCTTTGAGTCC 900
Qy 301 MetProAsnProGluGlyArgTrpThrPheGlyAlaSerCysValThrAlaCysProTrp 320

Db 901 ATGCCCAATCCGAGGCGCGGTATACATTGCGGCCAGCTGTGTGACTGCTGCCCTAC 960
Qy 321 AsnTrpLeuSerThrAspValGlySerAlaSerIleIleAsnPheGluLysLeuGlyAla 340
Db 961 AACTACCTTTCTACGACGCTGGGATCCGCTAGCATCATTAATTTTCGAGAAGTTGGCGCT 1020
Qy 341 GlyGlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeu 360
Db 1021 GGGGCGATGTCCACACAGGACCGCAGCTCATCTACAGGAGTGGCGGTGGGACCTG 1080
Qy 361 ThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGlu 380
Db 1081 ACATAGGCTGGAGCCCTCTGAAGAGAGGAGCCCGAGGTCTCCACTGGCACCTCCGAA 1140
Qy 381 GlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGln 400
Db 1141 GGGGCTGGCTCCGATGTATTGATGGTGACCTGGGAATGGGGGAGCCCAAGGGGCTGCA 1200
Qy 401 SerLeuProThrHisAspProSerProLeuGlnArgTrpSerGluAspProThrValPro 420
Db 1201 AGCTCCCCACACATGACCCCGCCCTCTACAGGGTACAGTGAGGACCCACAGTACCC 1260
Qy 421 LeuProSerGluThrAspGlyTrpValAlaProLeuThrCysSerProGlnProGluTrp 440
Db 1261 CTGCCCTCTGAGACTGATGGCTAGCTTGGCCCTGACCTGCAGCCCCCAGCCTGAATAT 1320
Qy 441 ValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAla 460
Db 1321 GTGAACCAAGCAGATGTTCGGCCCCCAGCCCCCTTCGCCCGAGAGGGCCCTCTGCCCTGCT 1380
Qy 461 AlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsnGly 480
Db 1381 GCCCGACCTGCTGGTGCCACTCTGGAAGGGGCCAAAGACTCTCTCCCGAGGAGAAATGGG 1440
Qy 481 ValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTrpLeuThrPro 500
Db 1441 GTCTCAAGACGTTTTTTTGGCTTTGGGGGTGCGCTGGAGAACCCCGAGTACTTGACACCC 1500
Qy 501 GlnGlyGlyAlaLaProGlnProHisProProAlaPheSerProAlaPheAspAsn 520
Db 1501 CAGGAGAGAGTGGCTCCCTCAGCCCCCAGCTCTCTGCTGCTTCCAGCCAGCCTTCAAGGG 1560
Qy 521 LeuTrpTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGly 540
Db 1561 CTCTATTACTGGGACCAAGGACCCACAGAGCGGGGGCTCCACCCAGCAGCCTTCAAGGG 1620
Qy 541 ThrProThrAlaGluAsnProGluTrpLeuGlyLeuAspValProAlaAlaHisHis 560
Db 1621 ACACCTACGGCAGAGAACCCAGAGTACCTGGGTCTGGAGCTGCCAGCGCGCCACATCAC 1680
Qy 561 HisHisHisHis 564
Db 1681 CATCACCATCAC 1692

RESULT 2

US-09-821-883-9
; Sequence 9, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laue, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821.883
; PRIOR FILING DATE: 2001-03-30
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 06:51:36 ; Search time 97.2259 Seconds
(without alignments)
3219.229 Million cell updates/sec

Title: US-09-821-883-3

Perfect score: 3062

Sequence: 1 MRAAPLLARAASLSGLFLF.....ENPEYLGVDVPAAHHHHHH 564

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO spool/US09821883/runat_09092004_105128_6657/app_query.fasta_1.3100
-DB=Issued_Patents_NA -QMT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOFL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09821883 @CGN 1.1 215 @runat_09092004_105128_6657 -NCPU=6 -ICPU=3
-NO WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTCUS COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2427.5	79.3	4530	1 US-08-229-515A-9	Sequence 9, Appli
2	2427.5	79.3	4530	1 US-08-645-865-9	Sequence 9, Appli
3	2427.5	79.3	4530	4 US-09-167-322-4	Sequence 4, Appli
4	2427.5	79.3	4530	4 US-09-527-487-1	Sequence 1, Appli
5	2427.5	79.3	4530	4 US-09-877-177A-11	Sequence 11, Appli
6	2425	79.2	9274	4 US-09-811-115-1	Sequence 1, Appli
7	2422.5	79.1	4473	2 US-09-048-804-1	Sequence 1, Appli
8	2422.5	79.1	4473	3 US-09-056-105-26	Sequence 26, Appli
9	2422.5	79.1	4473	4 US-09-663-834A-3	Sequence 3, Appli
10	2422	79.1	3768	4 US-09-811-115-2	Sequence 2, Appli
11	2417	78.9	3768	2 US-08-625-101-1	Sequence 1, Appli
12	2417	78.9	3768	2 US-08-356-786-1	Sequence 1, Appli

13	1968	64.3	3955	1 US-08-229-515A-14	Sequence 14, Appli
14	1968	64.3	3955	1 US-08-645-865-14	Sequence 14, Appli
15	1641.5	53.6	2385	1 US-09-146-283-3	Sequence 3, Appli
16	1641.5	53.6	2385	3 US-08-579-823A-3	Sequence 3, Appli
17	1641.5	53.6	2385	3 US-09-344-195-3	Sequence 3, Appli
18	1626.5	53.1	1872	3 US-08-422-108-2	Sequence 2, Appli
19	1626.5	53.1	1872	4 US-08-422-734-2	Sequence 2, Appli
20	732	23.9	1868	1 US-08-658-893B-1	Sequence 1, Appli
21	732	23.9	1868	4 US-09-676-610B-26	Sequence 26, Appli
22	731.5	23.9	1593	4 US-09-676-610B-25	Sequence 25, Appli
23	731.5	23.9	5532	2 US-08-475-035-3	Sequence 3, Appli
24	731.5	23.9	5532	4 US-09-676-610B-17	Sequence 17, Appli
25	723.5	23.6	5484	3 US-09-632-580A-3	Sequence 3, Appli
26	723.5	23.6	5551	1 US-08-484-438-1	Sequence 1, Appli
27	723.5	23.6	5551	1 US-08-484-438-3	Sequence 3, Appli
28	711	23.2	4905	1 US-07-978-895-3	Sequence 3, Appli
29	711	23.2	4905	1 US-08-473-119-3	Sequence 3, Appli
30	711	23.2	4905	2 US-08-475-352-3	Sequence 3, Appli
31	711	23.2	4905	4 US-09-170-689-3	Sequence 3, Appli
32	711	23.2	4975	3 US-09-630-706-3	Sequence 3, Appli
33	703	23.0	1958	4 US-09-570-454-1	Sequence 1, Appli
34	703	23.0	1958	4 US-09-867-521-1	Sequence 1, Appli
35	624.5	20.4	4545	6 5183884-3	Patent No. 5183884
36	493	16.1	322	1 US-08-421-356-1	Sequence 1, Appli
37	493	16.1	322	4 US-09-046-783-1	Sequence 1, Appli
38	259	8.5	4149	2 US-08-737-715-1	Sequence 1, Appli
39	259	8.5	4723	4 US-09-023-655-1137	Sequence 1137, Ap
40	255.5	8.3	2533	3 US-09-149-922-5	Sequence 5, Appli
41	242	7.9	4975	2 US-08-249-687C-1	Sequence 1, Appli
42	242	7.9	4989	2 US-08-666-392A-3	Sequence 3, Appli
43	242	7.9	4989	3 US-08-625-819-1	Sequence 1, Appli
44	242	7.9	4989	3 US-08-755-558-4	Sequence 4, Appli
45	242	7.9	4989	3 US-08-746-559A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-08-229-515A-9
; Sequence 9, Application US/08229515A
; Patent No. 5518885
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESS: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,515A
; FILING DATE: 19 Apr 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs

! TYPE: nucleic acid
! STRANDEDNESS: single
! TOPOLOGY: linear
US-08-229-515A-9

Alignment Scores:

Pred. No.: 1.86e-189 Length: 4530
Score: 2427.50 Matches: 522
Percent Similarity: 41.61% Conservative: 9
Best Local Similarity: 40.91% Mismatches: 20
Query Match: 79.28% Indels: 725
DB: 1 Gaps: 4

US-09-821-883-3 (1-564) x US-08-229-515A-9 (1-4530)

```
QY 2 ArgAlaLaProLeuLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17
DB 94 CGCGCCCTCCAGCGCGGTCCAGCGCGGCATGGGGCGGAGCGCGAGTGAGCACCATG 153
QY 18 PheLeuPheLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArg 37
DB 154 GAGCTGGCGGCTTGTGCGGTGG-----GGGCTCCTCTCTGCGCCCTCTTGCCCCC 204
QY 38 GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57
DB 205 GGAGCGCGAGCACCCAAAGTGTGCACCGGCACAGACATGAAGCTGGGCTCCCTGCCAGT 264
QY 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGly 77
DB 265 CCCGAGACCCACCTGGACATGCTCGCCACCTCTACACAGGGCTGCCAGGTGGTGCAGGA 324
QY 78 AsnLeuGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97
DB 325 AACCTGGAACTCACTACTGCGCCCAATGCCAGCCTGTCTTCTCCAGAGATATCCAG 384
QY 98 GluValGlnGlyTyrValLeuIleAlaHisGlnValArgGlnValProLeuGlnArg 117
DB 385 GAGGTGCAGGGCTACGTCTCATGCTCACAAACCAAGTGAGGCAGGTCCCACTGCAGAGG 444
QY 118 LeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAsp 137
DB 445 CTGCGGATTTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGGTGTAGAC 504
QY 138 AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157
DB 505 AATGAGACCCGCTGAAACATACACCCCTGTACAGGGGCTCCCCAGGAGGCTGCGG 564
QY 158 GluLeuGlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsn 177
DB 565 GAGCTGCAGCTTCGAAGCCTCACAGAGATCTTGAAGGAGGGGTCTTGTATCCAGCGAAC 624
QY 178 ProGlnLeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGln 197
DB 625 CCCAGCTCTGTACACAGACACGATTTTGTGGAAGGACATCTTCCACAAAGAACACAC 684
QY 198 LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217
DB 685 CTGGCTCTCACACTGATAGACACCAACCGCTCTCGGGCTGCGACCCCTGTTCTCCGATG 744
QY 218 CysLysGlySerArgCysTrpGlyGluSerGluAspCysGlnSerLeuThrArgThr 237
DB 745 TGTAAGGCTCCCGCTGTGGGGAGAGATTCTGAGGATTGTGAGAGCTTGCAGCGCACT 804
QY 238 ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysHisGlu 257
DB 805 GTCTGTGCGGTGGTGTGCGCGTGCAGGGGCCACTGCGCCACTGACTGCTGCCATGAG 864
QY 258 GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277
DB 865 CAGTGTGCTGCGGTGACCGGCCCCCAAGCACCTCTGACTGCTGGCTGCTGCTCCACTTC 924
QY 278 AsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThr 297
```

```
DB 925 AACCAAGAGTGTGAGCTGCTGAGCTGCCAGCCCTGGTCACTCAACAACACAGACACG 984
QY 298 PheGluSerMetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAla 317
DB 985 TTTGAGTCCATGCCCAATCCCGAGGCGGTATACATTCGGCGCCAGCTGTGACTGCC 1044
QY 318 CysProTyrAsnTyrLeuSerThrAspValGlySerAlaSerIleIle----- 333
DB 1045 TGTCCCTACAACTACCTTTCTACGAGCGTGGATCTGTCACCCCTCGTCTGCCCTGCAC 1104
QY 333 ----- 333
DB 1105 AACCAAGAGGTGCACAGCAGAGATGGAACACAGCGGTGTGAGAGTGCAGCAAGCCCTGT 1164
QY 333 ----- 333
DB 1165 GCCCGAGTGTCTATGGTCTGGGCATGGAGCACTTGGAGAGGTGAGGGCAGTTACCAAGT 1224
QY 333 ----- 333
DB 1225 GCCAATATCCAGAGATTGCTGGCTGCAAGAAGATCTTTGGAGGCTTGGCATTTCTGCGG 1284
QY 333 ----- 333
DB 1285 GAGAGCTTTGATGGGAGCCAGCCTCCAAACACTGCCCGCTCCAGCAGAGAGCAGCTCCAA 1344
QY 333 ----- 333
DB 1345 GTGTTTGAGACTCTGGAAGAGATCACAGGTTACCTATATACATCTCAGCATGCGCGGACAGC 1404
QY 333 ----- 333
DB 1405 CTGCTGACCTCAGGCTTTCCAGAACCTGCAAGTAATCCGGGAGCAATTCGCACAAT 1464
QY 333 ----- 333
DB 1465 GCGCCTACTCTGCTGACCCCTGCAAGGGCTGGGCATCAGCTGGCTGGGCTGCGCTCACTG 1524
QY 333 ----- 333
DB 1525 AGGGAACCTGGGCACTGGAGTGGCCCTCATCCACCATAACACCCACCTCTGCTTGTGTCAC 1584
QY 333 ----- 333
DB 1585 ACGGTGCCCTGGGACAGCTCTTTCGGAAACCCGACCAAGCTCTGCTCCACACTGCCAAC 1644
QY 333 ----- 333
DB 1645 CGGCCAGAGGACGAGTGTGTGGGCGAGGGCTGGCCCTGCCACCAGCTGTGCGCCCGAGGG 1704
QY 333 ----- 333
DB 1705 CACTGCTGGGTTCAGGGGCCACCCAGTGTGTCAACTGACCCAGTTCTTCGGGGGCCAG 1764
QY 333 ----- 333
DB 1765 GAGTGGTGGAGGAATGCGGAGTACTGCGGGGCTCCCGAGGAGTATGTAATGCCAGG 1824
QY 333 ----- 333
DB 1825 CACTGTTTTCGCTGCCACCTGAGTGTGAGCCCGAGGAGTGGCTCAGTGACCTGTTTGG 1884
QY 333 ----- 333
DB 1885 CCGAGAGGTGACAGTGTGTGGCCCTGTGCCCACTATTAAGGACCCCTCCCTCTGCGTGGCC 1944
QY 333 ----- 333
DB 1945 CGCTGCCCGAGCGTGTGAACCTGACCTCTCTACATGCCCACTGTGGAAGTTTCCAGAT 2004
QY 333 ----- 333
DB 2005 GAGGAGGGCGCATGCCAGCCTTGCCCCCATCAACTGCAACCCACTCCTGTGTGGACCTGGAT 2064
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:04:16 ; Search time 555.448 Seconds
(without alignments)
4313.608 Million cell updates/sec

Title: US-09-821-883-3

Perfect score: 3062

Sequence: 1 MRAAPLLARASLSGLFLF.....ENPEYGLDVPAAHHHHH 564

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/sgn2_1/USPTO spo01/US09821883/runat_09092004_105126_6623/app_query.fasta_1.3100
-DB=N Geneseq_29Jan04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blowum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09821883 @CGN 1.1.1646 @runat_09092004_105126_6623 -NCPU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_29Jan04:.*
1: geneseqn1980s:.*
2: geneseqn1990s:.*
3: geneseqn2000s:.*
4: geneseqn2001as:.*
5: geneseqn2001bs:.*
6: geneseqn2002a:.*
7: geneseqn2003as:.*
8: geneseqn2003bs:.*
9: geneseqn2003cs:.*
10: geneseqn2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3062	100.0	1692	4 AAD21566	Aad21566 Human HER
2	3014	98.4	2091	4 AAD21567	Aad21567 Human HER
3	3005.5	98.2	1665	4 AAD21564	Aad21564 Human HER
4	2957.5	96.6	2070	4 AAD21565	Aad21565 Human HER
5	2427.5	79.3	4530	2 AAT01585	Aat01585 Her-2/neu
6	2427.5	79.3	4530	2 AAT71253	Aat71253 Human HER
7	2427.5	79.3	4530	3 AA260815	AA260815 Nucleotid
8	2427.5	79.3	4530	4 AAD19731	Aad19731 Human tyr

9	2427.5	79.3	4530	6 ABN85585	Abn85585 Human HER
10	2427.5	79.3	4530	6 AB235012	Ab235012 Human gen
11	2427.5	79.3	4530	6 ABV94128	Abv94128 Breast ca
12	2427.5	79.3	4530	6 ABK83918	Abk83918 Human cDN
13	2427.5	79.3	4530	7 ACC50139	Acc50139 Breast ca
14	2427.5	79.3	4530	7 ABQ83856	Abq83856 Human Her
15	2427.5	79.3	4530	8 AAD58073	Aad58073 Human C-e
16	2427.5	79.3	4530	9 ADC09594	Adc09594 Her2/Neu
17	2425	79.2	9274	6 AAD43934	Aad43934 HER-2 tra
18	2425	79.2	9274	6 ABK14057	Abk14057 Human HER
19	2422.5	79.1	4472	3 AAA14812	Aaa14812 cDNA enco
20	2422.5	79.1	4473	2 ABQ76220	Abq76220 Human tum
21	2422.5	79.1	4473	2 AAZ31071	Aaz31071 HER-2 nuc
22	2422.5	79.1	4473	6 AB234969	Ab234969 Human gen
23	2422.5	79.1	4473	6 AAD38904	Aad38904 Human Her
24	2422.5	79.1	4473	7 ACC69999	Acc69999 Human C-e
25	2422.5	79.1	4473	9 ADC35148	Adc35148 Human bre
26	2422.5	79.1	4473	9 ADD25483	Add25483 Binding d
27	2422	79.1	3765	9 ADB67620	Adb67620 Human epi
28	2422	79.1	3768	3 AAA09455	Aaa09455 Human her
29	2422	79.1	3768	6 ABV78168	Abv78168 Human ERB
30	2422	79.1	3768	6 AB235744	Ab235744 Human ERB
31	2422	79.1	3768	6 AAD43935	Aad43935 Human HER
32	2422	79.1	3768	6 AEX09987	Abx09987 Human ERB
33	2422	79.1	3768	6 AAD43986	Aad43986 Human Her
34	2422	79.1	3768	6 ABK14058	Abk14058 Human HER
35	2422	79.1	3768	6 ABL91709	Ab191709 Human pol
36	2422	79.1	3768	7 ACC57649	Acc57649 Human pro
37	2418	79.0	9274	4 AAF24297	Aaf24297 HER2 tran
38	2417	78.9	3768	2 AAT40739	Aat40739 HER-2/neu
39	2417	78.9	3768	2 AAX01912	Aax01912 Human HER
40	2417	78.9	3768	5 AAD23392	Aad23392 Human HER
41	2417	78.9	3768	6 AAD32743	Aad32743 Human Her
42	2417	78.9	3768	6 ABK10730	Abk10730 Human Her
43	2417	78.9	3768	6 ABA92250	Ab92250 Human Her
44	2402.5	78.5	4606	9 ADB47370	Adb47370 Human cDN
45	2401	78.4	4299	2 AAQ46083	Aaq46083 Sequence

ALIGNMENTS

RESULT 1

AAD21566

ID AAD21566 standard; DNA; 1692 BP.

XX AC AAD21566;

XX DT 28-JAN-2002 (first entry)

XX DE Human HER500 fusion DNA construct comprising OVA-derived octapeptide.

XX KW Immunostimulatory fusion protein; IFP; antigen component; therapy;
KW immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human; OVA;
KW HER-2 protein; ovalbumin-derived octapeptide; HER500 fusion DNA; ds.

XX OS Homo sapiens.

OS Unidentified.

OS Chimeric.

XX FH Key

XX CDS

XX Location/Qualifiers

XX 1..1692

XX /*tag= a

XX /product= "Human HER500 fusion protein construct

XX comprising human PAP signal sequence, mature PAP protein,

XX an Ala Arg linker, human HER-2 signal sequence, mature

XX HER-2 membrane distal extracellular domain, an Ala

XX linker, an ovalbumin (OVA)-derived immunodominant

XX octapeptide, HER-2 membrane distal intracellular domain

XX and a C-terminal tag"

XX /note= "CDS does not include stop codon"

```

FT      /partial
XX      W0200174855-A2.
XX      11-OCT-2001.
XX      30-MAR-2001; 2001WO-US010515.
XX      30-MAR-2000; 2000US-0193504P.
XX      (DEND-) DENDREON CORP.
XX      Laus R, Vidovic D, Graddis T;
XX      WPI; 2001-662965/76.
XX      P-PSDB; ABE13110.
XX      An immunostimulatory fusion protein comprising the intracellular domain
XX      of HER-2 and an antigen elicits an immune response to the antigen and is
XX      useful for the treatment of associated cancer associated.
XX      Example 3; Page 28; 59pp; English.
XX      The invention relates to immunostimulatory fusion proteins (IFP) and
XX      nucleic acid molecules encoding such proteins. The IFPs comprise a
XX      polypeptide antigen component and an immunostimulatory component derived
XX      from the intracellular domain of HER-2 protein which is effective to
XX      elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
XX      immune response to the antigen. IFP or superactivated dendritic cells are
XX      used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
XX      associated with a particularly antigen. The present sequence is HER500
XX      fusion DNA construct which comprises DNA molecules encoding human PAP
XX      signal sequence, mature PAP protein, an Ala Arg linker, human HER-2
XX      signal sequence, mature HER-2 membrane distal extracellular domain, an
XX      Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide, HER-2
XX      membrane distal intracellular domain and a C-terminal tag
XX      SQ      Sequence 1692 BP; 338 A; 573 C; 470 G; 311 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      3.3e-146      Length:      1692
Score:          3062.00      Matches:      564
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              4          Gaps:      0

US-09-821-883-3 (1-564) x AAD21566 (1-1692)
QY      1 MetArgAlaProLeuLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPhe 20
Db      1 ATGAGAGTGACCCCTCTCTGCGCAGGCGCAGCAGCCTTAGCCTTGGCTTCTTGT 60
QY      21 LeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAla 40
Db      61 CTGCTTTTTTCTGCTAGACCGAAGTGCTACTAGCCAAAGGAGTTGGCGCGGGCGCG 120
QY      41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
Db      121 TCGACCCCAAGTGTGCACCGGCACACATGAGCTGCGGCTCCCTGCCAGTCCCGAGACC 180
QY      61 HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeuGlu 80
Db      181 CACCTGGACATGCTCCCGCACCTCTACAGGGCTCCAGGTGGTGCAGGGAAACCTGGAA 240
QY      81 LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnValGln 100
Db      241 CTCACCTACCTGCCCAACAAATGCCAGCCTGCTCTTCTCTGCAGATATCCAGAGGTGCAG 300
QY      101 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
Db      301 GGTCACGTGCTATCGCTCACAAACCAAGTAGGCGAGGTCCTCCACTGCAGAGGCTGCGGATT 360

```

```

QY      121 ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 140
Db      361 GTGCGAGGCACCCAGCTCTTTTGGAGACAACATATGCTGGCCGTGTAGACAATGGAGAC 420
QY      141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgLeuGln 160
Db      421 CCGCTGAACAATACACCCCTGTACAGGGGCCCTCCCAAGAGAGGCTGCGGAGCTGCAG 480
QY      161 LeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGlnLeu 180
Db      481 CTTCGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAAACCCCGAGCTC 540
QY      181 CysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200
Db      541 TGCTACCAGGACACGATTTTGTGAAGACATCTTCCACAAGAAACAACACCTGGCTCTC 600
QY      201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220
Db      601 ACATGTATAGACACCAACCGCTCTCGGGCCCTGCCACCCCTGTTCTCCGATGTGAAGGC 660
QY      221 SerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAla 240
Db      661 TCCCGCTGCTGGCGAGAGAGTTCTGAGATTGTTCAGAGCCTGACGCGCACTGTCTGTGCC 720
QY      241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCysAla 260
Db      721 GGTGGCTGTGCCCTGTCAGAGGGCCACTGCCCACTGACTGCTGCATGACGAGCTGTCT 780
QY      261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
Db      781 GCCGGCTGCACGGGCCCAAGCACTCTGACTGCTTGGCTGCTCCCTCCACTTCAACCAAGT 840
QY      281 GlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSer 300
Db      841 GGCATCTGTGAGCTGCATGCCAGCCCTGCTCACTACAAACACAGACACGTTTGAGTCC 900
QY      301 MetProLeuProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr 320
Db      901 ATGCCCAATCCGAGGGCCGCTATACATTCGCGCCAGCTGTGTGACTGCTGTCCCTTAC 960
QY      321 AsnTyrLeuSerThrAspValGlySerAlaSerIleLeuAsnPheGluLysLeuGlyAla 340
Db      961 AACTACCTTTCTACGAGCTGGGATCCCTAGCATCATTAATTTTCGAGAAGTTTGGCCCT 1020
QY      341 GlyGlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeu 360
Db      1021 GGGGCGATGGTCCACCACAGCCAGCCAGCTCATCTACAGAGTGGCGGTGGGACCTG 1080
QY      361 ThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGlu 380
Db      1081 ACATTAGGGCTGGAGCCCTCTGAAGAGAGAGGCCCCCAAGTCTCCACTGGCAACCTCCGAA 1140
QY      381 GlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGln 400
Db      1141 GGGGCTGCTCCGATGATTTGATGGTACCTGGGAATGGGGGCACCAAGGGGTGGAA 1200
QY      401 SerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValPro 420
Db      1201 AGCTCCCAACACATGACCCAGCCCTCTACAGCGGTACAGTGAAGACCCCAACAGTACC 1260
QY      421 LeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyr 440
Db      1261 CTGCGCTCTGAGACTGATGGCTAGCTTGGCCCTCTGACTGTCAGCGCCCGAGCTGAATAT 1320
QY      441 ValAsnGlnProAspValArgProGlnProProProProArgGluGlyProLeuProAla 460
Db      1321 GTGAACACAGCCAGATGTTTGGGCCCCAGCCCTTTCGCCCCAGAGAGGGCCCTCTGCTGCT 1380
QY      461 AlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsnGly 480
Db      1381 GCCCGACCTGCTGGTGGCCACTCTGGAAGGGCCCAAGACTCTCTCCCGAGGGAAGATGGG 1440
QY      481 ValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrPro 500

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:41:11 ; Search time 6042.41 Seconds
(without alignments)
4045.651 Million cell updates/sec

Title: US-09-821-883-3
Perfect score: 3062
Sequence: 1 MRAAPLLIARAASLSGLFLF.....ENPEYLGLDVPAAHHHHHH 564

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.apool/US09821883/runat_09092004.105127.6631/app.query.fasta_1.3100
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATH=human40.cdi -LIST=45
-DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09821883 @CEN 1.1 16795 @runat_09092004.105127.6631 -NCPU=6 -ICPU=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vi:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rod:*
36: em.htg.mam:*
37: em.htg.vrt:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3062	100.0	1692	6	AX268287 Sequence
2	3014	98.4	2091	6	AX268288 Sequence
3	3005.5	98.2	1665	6	AX268285 Sequence
4	2957.5	96.6	2070	6	AX268286 Sequence
5	2427.5	79.3	4530	6	I21124 Sequence 9
6	2427.5	79.3	4530	6	I59745 Sequence 9
7	2427.5	79.3	4530	6	AR202597 Sequence
8	2427.5	79.3	4530	6	AR283481 Sequence
9	2427.5	79.3	4530	6	AR344811 Sequence
10	2427.5	79.3	4530	6	AX282577 Sequence
11	2427.5	79.3	4530	6	AX587649 Sequence
12	2427.5	79.3	4530	6	AX644071 Sequence
13	2427.5	79.3	4530	6	AX771418 Sequence
14	2427.5	79.3	4530	6	BD005474 Cellular
15	2427.5	79.3	4530	9	HUMHER2A
16	2425	79.2	9274	6	AR409602 Sequence
17	2425	79.2	9274	6	AX060703 Sequence
18	2422.5	79.1	4473	6	AR080259 Sequence
19	2422.5	79.1	4473	6	AR167390 Sequence
20	2422.5	79.1	4473	6	AR392088 Sequence
21	2422.5	79.1	4473	9	HSERB2R
22	2422	79.1	3768	6	AR409603 Sequence
23	2422	79.1	3768	6	AX060704 Sequence
24	2422	79.1	3768	6	AX467229 Sequence
25	2422	79.1	3768	6	AX481438 Sequence
26	2422	79.1	3768	6	BD224136 Novel met
27	2417	78.9	3768	6	AR034479 Sequence
28	2417	78.9	3768	6	BD267514 HER-2/neu
29	2417	78.9	3768	6	AX201817 Sequence
30	2417	78.9	3768	6	AX380923 Sequence
31	2417	78.9	3768	6	AX384604 Sequence
32	2417	78.9	3768	6	AX465456 Sequence
33	2346	76.6	3678	6	AX505114 Sequence
34	2142.5	70.0	2763	6	AX380942 Sequence
35	2142.5	70.0	2781	6	AX380944 Sequence
36	2136	69.8	3780	4	AB008451 Canis fam
37	1990.5	65.0	4506	10	AK129487 Mus muscu
38	1977	64.6	4694	10	BC046811 Mus muscu
39	1977	64.6	4695	10	BC053078 Mus muscu
40	1974.5	64.5	3771	6	BD267516 HER-2/neu
41	1974.5	64.5	3771	6	AX189662 Sequence
42	1974.5	64.5	3771	6	AX380925 Sequence
43	1969.5	64.3	4062	10	HAMNEU
44	1968.5	64.3	4727	10	AY116182 Rattus no
45	1968	64.3	3955	6	BD267515 HER-2/neu

ALIGNMENTS

AX268287
LOCUS AX268287 1692 bp DNA linear PAT 29-OCT-2001
DEFINITION Sequence 8 from Patent WO0174855.
ACCESSION AX268287
VERSION AX268287.1 GI:16541539
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE
1 Laus, R., Vidovic, D. and Graddis, T.
AUTHORS Compositions and methods for dendritic cell-based immunotherapy
TITLE Patent: WO 0174855-A 8 11-OCT-2001;
JOURNAL DENDREON CORPORATION (US)
FEATURES
source Location/Qualifiers
1..1692
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="HER500* construct"

ORIGIN
Alignment Scores:
Pred. No.: 7e-119 Length: 1692
Score: 3062.00 Matches: 564
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-821-883-3 (1-564) x AX268287 (1-1692)

Qy 1 MetArgAlaProLeuLeuLeuAlaArgAlaSerLeuSerLeuGlyPheLeuPhe 20
Db 1 ATGAGAGTGCACCCCTCTCTGCGCCAGGCGAGCAGCTTAGCCTTTGGCTTCTGTTT 60
Qy 21 LeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyValAla 40
Db 61 CTGCTTTTCTGGCTAGACCGAAGTGTACTAGCCAGAGGATTGGCCGCGGGCCGCG 120
Qy 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
Db 121 TCGACCCCAAGTGTGCACCGGCACACATGAAGCTGGCGCTCCCTGCCAGTCCCGAGACC 180
Qy 61 HisLeuAspMetLeuArgHisLeuTrpGlnGlyCysGlnValValGlnGlyAsnLeuGlu 80
Db 181 CACCTGGACATGCTCCCGCACCTCTTACCAGGGCTGCCAGGTGGTGCAGGGAACCTGGAA 240
Qy 81 LeuThrTrpLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGln 100
Db 241 CTCACCTACCTGCCCCACCAATGCCAGCTGTCTTCTTCCAGAGATATCCAGAGGTGCAG 300
Qy 101 GlyTrpValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgile 120
Db 301 GGCTACGTGCTCATCGCTCACCAAGTGAAGGAGGTCCACATGCAGAGCTCGCGATT 360
Qy 121 ValArgGlyThrGlnLeuPheGluAspAsnTrpAlaLeuAlaValLeuAspAsnGlyAsp 140
Db 361 GTGGAGGACCCAGCTCTTTGAGGACAACTATGCTTGGCGCTGGCTAGACAAATGGAGAC 420
Qy 141 ProLeuAsnAsnThrProValThrGlyValAserProGlyGlyLeuArgGluLeuGln 160
Db 421 CCGCTGAACAATATACCCCTGTACAGGGGCTCCCGAGGAGGCTCGGGAGCTGCAG 480
Qy 161 LeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsnProGlnLeu 180
Db 481 CTTCGAACCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAAACCCCGAGCTC 540
Qy 181 CysTrpGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200
Db 541 TGCTACCAAGGACACGATTTTGTGAAGGACATCTTCCACAAGAAACACAGCTGGCTCTC 600
Qy 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220

601 ACATGTATAGACACCAACCGCTCTCGGGCCCTGCGCACCCCTGTTCTCCGATGTGTAAGGGC 660
221 SerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAla 240
661 TCCGCTGCTGGGAGAGATTCTGAGGATTTGTGAGAGCTGAGCGGACACTGTCTGTGCC 720
241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCysAla 260
721 GGTGGCTGTGCCCTGCAAGGGGCCACTGCGCACTACTGCTGCTCATGAGCACTGTGCT 780
261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
781 GCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
281 GlyIleCysGluLeuHisCysProAlaLeuValThrTrpAsnThrAspThrPheGluSer 300
841 GGCATCTGTGAGCTGCATGCCACCGCTGCTACCTACCAACACAGACACAGCTTTGAGTCC 900
301 MetProAsnProGluGlyArgTrpThrPheGlyAlaSerCysValThrAlaCysProTrp 320
901 ATGCCCAATCCGAGGGCCGGTATACATTGGCGCCAGCTGTGACTGCTGCTGCTGCTGCT 960
321 AsnTrpLeuSerThrAspValGlySerAlaSerIleLeuPheGluLysLeuGlyAla 340
961 AACTACCTTTCTACGGAGCTGGGATCCGCTAGCATCATTAATTCGAGAGTTCGGCGCT 1020
341 GlyGlyMetValHisHisArgHisArgSerSerSerThrArgSerGlyGlyGlyAspLeu 360
1021 GGGGCGATGGTCCACACAGCAGCCGAGCTCATCTACAGGAGTGGCGGTGGGAGCCTG 1080
361 ThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGlu 380
1081 AACTAGGCTGGAGCCCTCTGAAGAGAGGCCCCCGAGGTCTCCACTGGCACCCCTCCGAA 1140
381 GlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGln 400
1141 GGGGCTGGCTCCGATGATTTGATGTTGACTGGGAATGGGGGAGCAGCAAGGGGCTGCAA 1200
401 SerLeuProThrHisAspProSerProLeuGlnArgTrpSerGluAspProThrValPro 420
1201 AGCTTCCCAACACATGATGCCCGCTCTACAGCGGTACAGTGGAGACCCCGACAGTACCC 1260
421 LeuProSerGluThrAspGlyTrpValAlaProLeuThrCysSerProGlnProGluTrp 440
1261 CTGGCTCTGAGACTGATGGCTAGCTGTTGCCCTGACCTGACGCCCCCGACCTGGAATAT 1320
441 ValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAla 460
1321 GTGAACCGAGCAGATGTTGGGCCCGACGCCCTTCCGCCCGAGAGGGCCCTCTGCTGCT 1380
461 AlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsnGly 480
1381 GCCCGACCTGTGTGGTGCCTCTGGAAGGGGCCAAGACTCTCTCCCGAGGAAGAATGGG 1440
481 ValValLysAspValPheAlaPheGlyValAlaValGluAsnProGluTrpLeuThrPro 500
1441 GTCGTCAAGACGTTTTTGGCTTTGGGGTGGCTGGAGAACCCCGAGTACTTGNACACC 1500
501 GlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsn 520
1501 CAGGAGGAGCTGCCCTCAGCCCCACCTCTCTGCTTTCAGCCCCAGCCTTCGACAAC 1560
521 LeuTrpTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGly 540
1561 CTCTATTACTGGGACCGAGGCCACAGAGCGGGGGCTCCACCCAGCAGCCTTCAAGGG 1620
541 ThrProThrAlaGluAsnProGluTrpLeuGlyLeuAspValProAlaAlaHisHis 560
1621 ACACCTACCGCAGAGAACCCAGAGTACTGGGTCTGGNCGTGGCAGCGGCCGACATCAC 1680
561 HisHisHisHis 564
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:28:07 ; Search time 56.0399 Seconds
(without alignments)
3175.456 Million cell updates/sec

Title: US-09-821-883-3
Perfect score: 3062
Sequence: 1 MRAAPLLARAASLSGLFLP.....ENPEYGLDVPAAHHHHH 564

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2136	69.8	1259	6	018735
2	1862.5	64.1	1259	11	Q8K3F9
3	1612.5	52.7	419	4	Q9UK79
4	1380	45.1	711	11	Q80Y89
5	1014	33.1	881	11	Q8C0E7
6	966	31.5	367	11	Q8R2X1
7	908	29.6	165	4	Q14256
8	894.5	29.2	412	4	Q8WTV0
9	859	28.1	431	13	Q7SY19
10	751	24.5	527	13	Q90836
11	723	23.6	149	6	Q9BG66
12	717.5	23.4	643	11	Q9ERV6
13	717.5	23.4	655	11	Q9WVF5
14	717.5	23.4	1210	11	Q9EP98
15	713.5	23.3	1209	11	Q9GX70
16	708.5	23.1	1209	6	Q8MIL8

17	703	23.0	1191	13	Q7SZF7	Q7szf7 brachydanio
18	701.5	22.9	478	11	Q9ESE0	Q9ese0 rattus norv
19	697	22.8	331	4	Q9BUD7	Q9bud7 homo sapien
20	660	21.6	1165	13	Q9YH40	Q9yh40 xiphophorus
21	658	21.5	1305	13	Q8AW81	Q8aw81 brachydanio
22	635.5	20.8	599	13	Q9PSH2	Q9psh2 gallus gall
23	610	19.9	1328	13	P79754	P79754 fugu rubrip
24	559.5	18.3	1377	5	Q8MLW0	Q8mlw0 drosophila
25	543	17.7	1432	5	Q9BTH9	Q9bth9 anopheles g
26	540.5	17.7	1322	5	Q86NZ2	Q86nz2 drosophila
27	409.5	13.4	1137	13	Q9W6F6	Q9w6f6 gallus gall
28	400.5	13.1	150	6	Q9BG64	Q9bg64 oryctolagus
29	395	12.9	151	6	Q9BG65	Q9bg65 oryctolagus
30	377	12.3	1368	5	Q23821	Q23821 caenorhabdi
31	341	11.1	366	5	Q26569	Q26569 schistosoma
32	341	11.1	1717	5	Q26566	Q26566 schistosoma
33	331	10.8	334	5	Q26567	Q26567 schistosoma
34	331	10.8	342	5	Q26568	Q26568 schistosoma
35	321.5	10.5	1193	5	Q9YLX8	Q9ylx8 ephydatia f
36	305	10.0	1564	5	Q86MD7	Q86md7 echinococcu
37	290.5	9.5	1472	5	Q9USA8	Q9usa8 bombyx mori
38	289	9.4	1671	5	Q9NJV5	Q9njv5 biophalari
39	276.5	9.0	1749	5	Q8TOW6	Q8tow6 echinococcu
40	273.5	8.9	89	11	Q88459	Q88459 mus musculu
41	273.5	8.9	1418	13	Q8UW83	Q8uw83 paralichthy
42	272.5	8.9	1418	13	Q93457	Q93457 scophthalmu
43	271	8.9	2144	5	Q9VD94	Q9vd94 drosophila
44	263	8.6	1358	13	Q73798	Q73798 xenopus lae
45	262.5	8.6	1369	13	Q8UW86	Q8uw86 paralichthy

ALIGNMENTS

RESULT 1

ID	018735	PRELIMINARY;	PRT;	1259	AA.
AC	018735;				
DT	01-JAN-1998	(TrEMBLrel. 05, Created)			
DT	01-JAN-1998	(TrEMBLrel. 05, Last sequence update)			
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)			
DE	Erbb-2.				
OS	Canis familiaris (Dog).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.				
NC	NCBI_TaxID=9615;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Yokota H.;				
RT	"cDNA cloning of erbB-2 from canine mammary gland.";				
RL	Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AB008451; BAA23127.1; -				
DR	HSSP; P11362; 1FGK.				
DR	GO; GO:0016020; C:membrane; IEA.				
DR	GO; GO:0005524; F:ATP binding; IEA.				
DR	GO; GO:0005509; F:calcium ion binding; IEA.				
DR	GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.				
DR	GO; GO:0016740; F:transferase activity; IEA.				
DR	GO; GO:0004688; P:protein amino acid phosphorylation; IEA.				
DR	GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . . ; IEA.				
DR	InterPro; IPR002048; EF-hand.				
DR	InterPro; IPR006211; Furin-like.				
DR	InterPro; IPR000494; EGFR_L domain.				
DR	InterPro; IPR006212; Furin repeat.				
DR	InterPro; IPR009030; Grow_fac_recep.				
DR	InterPro; IPR000719; Prot_kinase.				
DR	InterPro; IPR001245; Tyr_kinase.				
DR	InterPro; IPR008266; Tyr_kinase_AS.				
DR	InterPro; IPR004019; YLP motif.				
DR	Pfam; PF00757; Furin-like; 1.				
DR	Pfam; PF00069; pkinase; 1.				
DR	Pfam; PF01030; Recep_L_domain; 2.				
DR	Pfam; PF02757; YLP; 2.				

DR	PRINTS; P000109; TYRKINASE.	
DR	ProDom; PD000001; Prot_kinase; 1.	
DR	SMART; SM00261; FU; 3.	
DR	SMART; SM00219; TyrKc; 1.	
DR	PROSITE; P500018; EF_HAND; 1.	
DR	PROSITE; P500107; PROTEIN_KINASE_DOM; 1.	
DR	PROSITE; P500111; PROTEIN_KINASE_DOM; 1.	
DR	PROSITE; P500109; PROTEIN_KINASE_TYR; 1.	
KW	ATP-binding; kinase; transferase; Tyrosine-protein kinase.	
SQ	SEQUENCE 1259 AA; 137989 MW; E37364D49C4ACD46 CRC64;	
Query Match 69.8%; Score 2136; DB 6; Length 1259;		
Best Local Similarity 36.9%; Pred. No. 5.6e-155;		
Matches 462; Conservative 22; Mismatches 47; Indels 722; Gaps 3;		
QY	25 WLDKRVLAKEALARGAASSTQVCTGDMKRLPASPETHLDMLRHLRYOGCOVQGNLELYL 84	
DB	6 WCRWGLLALLPSGAAGTQVCTGDMKRLPASPETHLDMLRHLRYOGCOVQGNLELYL 65	
QY	85 PTNLSLSELDIOEQVGVYLIHNOVROVPLQRLRIVRGTLFEDNYALAVLDNGDPLNN 144	
DB	66 PANASLSFLQDIQEQVGVYLIHNSQVRQIPLQRLRIVRGTLFEDNYALAVLDNGDPLEG 125	
QY	145 TTPVTGASPGGLRELQRLSLTEILKGGVLIQNPOLCYQDTILWKDIFHKNNQLALTLD 204	
DB	126 GIPAFGNAQGLRELQRLSLTEILKGGVLIQNSPOLCHQDTILWKDVFHKNNQLALTLD 185	
QY	205 TNRSRACHPCSPMKGSRGWGSSBDCQSLTRTVAGGCARCKGPLPTDCCHQCAGCT 264	
DB	186 TNRFSGPCSPACKDAHCWGASSGDCQSLTRTVAGGCARCKGQPTDCCHQCAGCT 245	
QY	265 GPKHSDCLAFHNSGICELHCPALVTNTDTFESMNPBGRTYFGASCVTACPNVYL 324	
DB	246 GPKHSDCLAFHNSGICELHCPALVTNTDTFESMNPBGRTYFGASCVTSCPNVYL 305	
QY	325 TDVGSASII----- 333	
DB	306 TDVGSCTLVCLPNNQEVTAEDGTQRCCKSKPCARVCYGLGMEHLREVRVAVTSANIQEP 365	
QY	334 ----- 333	
DB	366 GCKIFGSLAFPEFSGDGPASNTAPLQPEQLRVFEALBEITGYLYISAWPDSLNLVVF 425	
QY	334 ----- 333	
DB	426 QNLVRVGRVLDGAYSLTLQGLIGISWLGSLRLSRLGSLALIHNRALCFVHTVPWDL 485	
QY	334 ----- 333	
DB	486 FRNPQALLHSANRPEEBECVGBGLACYPCHAGHCWGPGTQCVCNSQFLRGOECVECRV 545	
QY	334 ----- 333	
DB	546 LQGLPREYKDYCLPCHSECOPONGSVTCFGSEADQCACAHYKDPFPCVACRPSGVKP 605	
QY	334 ----- 333	
DB	606 DLSFMPFIWFADEGTGCPQPCINCTHSCADLDEKGPABQASPVTSIIAAVVGILLAV 665	
QY	334 ----- 333	
DB	666 VGLVLGILIKRRQKIRKYMRLLQETLVEPLTPSGAMPNQAMQILKETELRKVKVL 725	
QY	334 ----- 333	
DB	726 GSGAFGVYKGIWIPDGENVKIPVAIKVLRENTSPKANKEILDEAYVWAGVSPVSRLL 785	
QY	334 ----- 333	
DB	786 GICLTSTVLTQMPYGCGLLDHVRHGRGLSGQDLLNWCVQIAKMSYLEDVRLVHRDL 845	
QY	334 ----- 333	

DB	846 AARNVLKSPNHVKITDFGLARLLDIDETEVHADGGKVPKIMWALESIPPRRFTHQSDVW 905	
QY	334 ----- 333	
DB	906 SYGVTVWELMTFGAKPYDIPAREIPDLLEKGERLPQPPICITDVTYMMVKCMWIDSECR 965	
QY	334 -NFEKL----- 338	
DB	966 PRFRELVAEFSEMRMDPQRFVVIQNEIDGPASPLDSTFVRSLEDDDDMGDLVDAEYLVP 1025	
QY	339 -----GAGMWHRRHRSSTRSGGDLTLGLSPSEEEAPRSLAPSEGAGSDVF 387	
DB	1026 QQGFPCPEPTFCAGCTAHRHRSSTRNGGGLTLGLSPSEEPKPSLAPSEGAGSDVF 1085	
QY	388 DGDLMGAAKGLQSLPDPQLQRYSEDPTVPLPSETDGVVAPLTCSPOPEYVNPQDVR 447	
DB	1086 DGDLMGAAKGLQSLPDPQLQRYSEDPTVPLPSETDGVVAPLTCSPOPEYVNPQVW 1145	
QY	448 PQPPSPREGPIPAARPAAGATLER-----AKTLSPGKNGVVKDVFVAFGGAVENPEYLTPOG 502	
DB	1146 PQPLALSGPLPPSPAGATLERPKTLSPKTLSPGKNGVVKDVFVAFGSAVENPEYLAERG 1205	
QY	503 GAAPQHPHPPAFSPAFDNLVYWDQPPPERGAPPSTFKGTPTAENPEYLGLDVP 555	
DB	1206 RAAQHPHPPAFSPAFDNLVYWDQPPSERGSPSTFEGTPTAENPEYLGLDVP 1258	
RESULT 2		
ID	Q8K3F9	
AC	Q8K3F9; PRELIMINARY; PRT; 1259 AA.	
DT	01-OCT-2002 (TrEMBLrel. 22, Created)	
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)	
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	
DE	Neu protooncoprotein.	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OX	NCBI_TaxID=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=BDIX;	
RA	Watson P.A., Kim K., Chen K.-S., Gould M.N.;	
RT	"Androgen-Dependent Mammary Carcinogenesis in Rats Transgenic for the	
RT	New Proto-Oncogene";	
RL	Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.	
DR	EMBL; AY116182; AAM50093.1; -	
DR	GO; GO:0016020; C:membrane; IEA.	
DR	GO; GO:0005524; F:ATP binding; IEA.	
DR	GO; GO:0005509; F:calcium ion binding; IEA.	
DR	GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.	
DR	GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.	
DR	GO; GO:0016740; F:transferase activity; IEA.	
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.	
DR	GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; IEA.	
DR	InterPro; IPR002048; EF-hand.	
DR	InterPro; IPR000494; EGFR_L domain.	
DR	InterPro; IPR006211; Furin-like.	
DR	InterPro; IPR009030; Grow_fac recep.	
DR	InterPro; IPR000719; Prot_kinase.	
DR	InterPro; IPR002290; Ser_thr_kinase.	
DR	InterPro; IPR001245; Tyr_kinase.	
DR	InterPro; IPR008266; Tyr_kinase_AS.	
DR	InterPro; IPR004019; YLP_motif.	
DR	Pfam; PF00757; Furin-like; 1.	
DR	Pfam; PF00069; kinase; 1.	
DR	Pfam; PF01030; Recep_L_domain; 2.	
DR	Pfam; PF02757; YLP; 2.	
DR	PRINTS; PR00109; TYRKINASE	
DR	ProDom; PD000001; Prot_kinase; 1.	
DR	SMART; SM00220; S_TKc; 1.	
DR	SMART; SM00219; TyrKc; 1.	
DR	PROSITE; PS00018; EF_HAND; 1.	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:20:54 ; Search time 10.3528 Seconds
(without alignments)
2836.691 Million cell updates/sec

Title: US-09-821-883-3

Perfect score: 3062

Sequence: 1 MRAAPLLARASLSGLFL.....ENPEYGLDVPAAHHHHH 564

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2417	78.9	1255	1 ERB2 HUMAN	P04626 homo sapien
2	1959	64.0	1257	1 ERB2 RAT	P06494 rattus norv
3	1946.5	63.6	1254	1 ERB2 MESAU	P06553 mesocricetu
4	752	24.6	703	1 EGFR CHICK	P13387 gallus gall
5	723.5	23.6	1308	1 ERB4 HUMAN	Q15303 homo sapien
6	719.5	23.5	1210	1 EGFR HUMAN	P00533 homo sapien
7	719.5	23.5	1308	1 ERB4 RAT	Q62356 rattus norv
8	717.5	23.4	1210	1 EGFR MOUSE	Q01279 mus musculu
9	697	22.8	1342	1 ERB3 HUMAN	P21860 homo sapien
10	668	21.8	1339	1 ERB3 RAT	Q62799 rattus norv
11	651	21.3	1167	1 XMRK XIPMA	P13388 xiphophorus
12	539.5	17.6	1426	1 EGFR DROME	P04412 drosophila
13	404.5	13.2	1367	1 LTR3 CAEL	P24348 caenorhabdi
14	318	10.4	1363	1 ILPR BRALA	Q02466 brachyosteo
15	301.5	9.8	245	1 ERB2 MOUSE	P70424 mus musculu
16	276.5	9.0	1477	1 HTK7 HYDAT	Q25197 hydra atten
17	271	8.9	2146	1 INSR DROME	P09208 drosophila
18	258	8.4	1382	1 INSR HUMAN	P06213 homo sapien
19	252.5	8.2	1372	1 INSR MOUSE	P15208 mus musculu
20	252	8.2	1607	1 MIPR LYMT	Q25410 lymanaea sca
21	251.5	8.2	1300	1 IRR MOUSE	Q9wt14 mus musculu
22	249	8.1	1383	1 INSR RAT	P15127 rattus norv
23	248.5	8.1	1297	1 IRR HUMAN	P14616 homo sapien
24	246.5	8.1	1300	1 IRR CAVPO	P14617 cavia porce
25	238.5	7.8	1367	1 IGRF HUMAN	P08069 homo sapien
26	236.5	7.7	1370	1 IGRF RAT	P24062 rattus norv
27	234.5	7.7	1373	1 IGRF MOUSE	Q60751 mus musculu
28	228.5	7.5	581	1 IRR RAT	Q64716 rattus norv
29	207	6.8	1390	1 INSR AEDAE	Q93105 aedes aegypt
30	168	5.5	386	1 PPAP HUMAN	P15309 homo sapien
31	155	5.1	5262	1 MLG2 HUMAN	Q14686 homo sapien
32	153.5	5.0	707	1 SFPO HUMAN	P23246 homo sapien
33	148.5	4.8	1321	1 IRS2 MOUSE	P81122 mus musculu

34	141.5	4.6	886	1	SM6B MOUSE	O54951 mus musculu
35	140	4.6	1696	1	PCK5 BRACL	Q9nj15 brachyosteo
36	139	4.5	634	1	ERBB ALV	P00534 avian leuko
37	139	4.5	2442	1	CBP HUMAN	Q92793 homo sapien
38	138	4.5	830	1	SREB HUMAN	Q14162 homo sapien
39	136	4.4	331	1	PRP1 HUMAN	P04280 homo sapien
40	135.5	4.4	553	1	ODO2 MYCTU	Q10381 mycobacteri
41	135	4.4	620	1	EXTN TOBAC	P13983 nicotiana t
42	134.5	4.4	863	1	MI13 HUMAN	Q8n3f8 homo sapien
43	132	4.3	833	1	SRC2 MOUSE	P59222 mus musculu
44	130.5	4.3	296	1	CC01 CAEL	P08124 caenorhabdi
45	130	4.2	1051	1	ULK1 MOUSE	O70405 mus musculu

ALIGNMENTS

RESULT 1
ERB2_HUMAN
ID ERB2_HUMAN STANDARD; PRT; 1255 AA.
AC P04626;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell
DE surface receptor HER2) (MLN 19).
GN ERB2 OR HER2 OR NGL OR NEU.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86118663; PubMed=3003577;
RA Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N.,
RA Saito T., Toyoshima K.;
RT "Similarity of protein encoded by the human c-erb-B-2 gene to
RT epidermal growth factor receptor.";
RL Nature 319:230-234(1986).
[2]
RN
RP SEQUENCE FROM N.A., AND VARIANT ALA-1170.
RX MEDLINE=86070181; PubMed=2999974;
RA Cousens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A.,
RA McGrath J., Seeburg P.H., Libermann T.A., Schlessinger J.,
RA Francke U., Levinson A., Ullrich A.;
RT "Tyrosine kinase receptor with extensive homology to EGF receptor
RT shares chromosomal location with neu oncogene.";
RL Science 230:1132-1139(1985).
[3]
RN
RP SEQUENCE FROM N.A., AND VARIANTS CYS-452; VAL-655 AND ALA-1170.
RX Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M.-W.,
RX Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RX Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RT Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
[4]
RN
RP SEQUENCE OF 737-1031 FROM N.A.
RX MEDLINE=86016729; PubMed=2995967;
RA Semba K., Kamata N., Toyoshima K., Yamamoto T.;
RT "A v-erbB-related protooncogene, c-erbB-2, is distinct from the
RT c-erbB-1/epidermal growth factor-receptor gene and is amplified in a
RT human salivary gland adenocarcinoma.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).
[5]
RN
RP VARIANTS VAL-654 AND VAL-655.
RX MEDLINE=93194196; PubMed=8095488;
RA Ehsani A., Low J., Wallace R.B., Wu A.M.;
RT "Characterization of a new allele of the human ERBB2 gene by allele-
RT specific competition hybridization.";
RL Genomics 15:426-429(1993).
CC
-!- FUNCTION: Essential component of a neurotrophin-receptor complex,
CC although neurotrophins do not interact with it alone. GF30 is a
CC potential ligand for this receptor. Not activated by EGF, TGF-

CC alpha and amphiregulin.
 CC -|- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -|- SUBUNIT: Heterodimer with each of the other ERBB receptors
 CC (Potential). Interacts with PRKCAP (By similarity).
 CC -|- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -|- PTM: Ligand-binding increases phosphorylation on tyrosine
 CC residues (By similarity).
 CC -|- POLYMORPHISM: There are four alleles due to the variations in
 CC positions 654 and 655. Allele B1 (ile-654/Val-655) has a frequency
 CC of 0.782; allele B2 (ile-654/Val-655) has a frequency of 0.206;
 CC allele B3 (Val-654/Val-655) has a frequency of 0.012.
 CC -|- SIMILARITY: Belongs to the EGF receptor family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M11767; AAA35808.1; -.
 CC EMBL; M11761; AAA35808.1; JOINED.
 CC EMBL; M11762; AAA35808.1; JOINED.
 CC EMBL; M11763; AAA35808.1; JOINED.
 CC EMBL; M11764; AAA35808.1; JOINED.
 CC EMBL; M11765; AAA35808.1; JOINED.
 CC EMBL; M11766; AAA35808.1; JOINED.
 CC EMBL; M11730; AAA75493.1; -.
 CC EMBL; M12036; AAA35978.1; -.
 CC EMBL; AX208911; AAQ18082.1; -.
 CC EMBL; X03363; CAA27060.1; -.
 CC FIR; A24571; A24571.
 CC PDB; 1N82; 18-FEB-03.
 CC PDB; 1QR1; 01-JAN-00.
 CC Genew; HGNC:3430; ERBB2.
 CC MIM; 164870; -.
 CC GO; GO:0005012; F.Receptor activity; TAS.
 CC GO; GO:0004716; F.Receptor signaling protein tyrosine kinase . . .; TAS.
 CC GO; GO:0008283; P.cell proliferation; TAS.
 CC GO; GO:0007167; P.enzyme linked receptor protein signaling pa. . .; TAS.
 CC GO; GO:0006470; P.protein amino acid dephosphorylation; TAS.
 CC GO; GO:0006468; P.protein amino acid phosphorylation; TAS.
 CC InterPro; IPR000494; EGFR_L domain.
 CC InterPro; IPR006211; Furin-like.
 CC InterPro; IPR006212; Furin repeat.
 CC InterPro; IPR009030; Grow fac recep.
 CC InterPro; IPR000719; Prot kinase.
 CC InterPro; IPR001245; Tyr_kinase.
 CC InterPro; IPR008266; Tyr_kinase_AS.
 CC InterPro; IPR004019; YLP motif.
 CC Pfam; PF00757; Furin-like; 1.
 CC Pfam; PF00069; pkinase; 1.
 CC Pfam; PF01030; Recep_L_domain; 2.
 CC Pfam; PF02757; YLP_2.
 CC PRINTS; PR00109; TYRKINASE.
 CC ProDom; PD000001; Prot_kinase; 1.
 CC SMART; SM00261; FU; 4.
 CC SMART; SM00219; TYRK; 1.
 CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 CC PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 CC Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
 CC Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 CC Polymorphism; 3D-structure.
 CC SIGNAL 1 21 POTENTIAL.
 CC CHAIN 22 1255 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.
 CC DOMAIN 22 652 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 653 675 POTENTIAL.
 CC DOMAIN 676 1255 CYTOPLASMIC (POTENTIAL).
 CC DOMAIN 720 987 PROTEIN KINASE.
 CC NP_BIND 726 734 ATP (BY SIMILARITY).
 CC -----

753 753 FT BINDING
 845 845 FT ACT_SITE
 195 195 FT DISULFID
 204 204 FT DISULFID
 139 139 FT DISULFID
 212 212 FT DISULFID
 227 227 FT DISULFID
 235 235 FT DISULFID
 244 244 FT DISULFID
 252 252 FT DISULFID
 264 264 FT DISULFID
 288 288 FT DISULFID
 299 299 FT DISULFID
 311 311 FT DISULFID
 331 331 FT DISULFID
 334 334 FT DISULFID
 511 511 FT DISULFID
 520 520 FT DISULFID
 535 535 FT DISULFID
 540 540 FT DISULFID
 560 560 FT DISULFID
 576 576 FT DISULFID
 584 584 FT DISULFID
 596 596 FT DISULFID
 600 600 FT DISULFID
 623 623 FT DISULFID
 634 634 FT DISULFID
 642 642 FT DISULFID
 1139 1139 FT MOD_RES
 1248 1248 FT MOD_RES
 68 68 FT CARBOHYD
 124 124 FT CARBOHYD
 187 187 FT CARBOHYD
 259 259 FT CARBOHYD
 530 530 FT CARBOHYD
 571 571 FT CARBOHYD
 629 629 FT CARBOHYD
 452 452 FT VARIANT
 654 654 FT VARIANT
 655 655 FT VARIANT
 1170 1170 FT VARIANT
 1255 1255 SQ SEQUENCE
 137909 137909 MW; 39E9FDA04DCF962 CRC64;
 78.9%; Score 2417; DB 1; Length 1255;
 41.2%; Pred No. 9.8e-138;
 4; Mismatches 6; Indels 718; Gaps 2;
 Query Match
 Best Local Similarity
 Matches 511; Conservative
 35 LARGAASQVCTGTDMLRLPASPTHLDMLRHLVGGQVQVQGNLELTYLPTNASSLFLQ 94
 16 LPPGAASQVCTGTDMLRLPASPTHLDMLRHLVGGQVQVQGNLELTYLPTNASSLFLQ 75
 95 DIQEVQGVYLIHNNQVQVPLQRLIRVGTQQLFEDNYALAVLDNGDPLNLTPTVTGASPG 154
 76 DIQEVQGVYLIHNNQVQVPLQRLIRVGTQQLFEDNYALAVLDNGDPLNLTPTVTGASPG 135
 155 GLRELQLRSLTEILKGGVLIQNPOLCYQDTTLWKDI PHKNQALATLTDNRSRACHPC 214
 136 GLRELQLRSLTEILKGGVLIQNPOLCYQDTTLWKDI PHKNQALATLTDNRSRACHPC 195
 215 SPMKSGRCWGESSEDCQSLTRTVAGGCARCKGLPTDCCHQCAAGCTGPRHSDCLAC 274
 196 SPMKSGRCWGESSEDCQSLTRTVAGGCARCKGLPTDCCHQCAAGCTGPRHSDCLAC 255
 275 LHFNHSGICELCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYLTVDVGSASII- 333
 256 LHFNHSGICELCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYLTVDVGSCTIVC 315
 334 334 QY ----- 333
 316 PLHNOEVTAEADCTQRCCKSPCARVCYGLGMEHLREVRAVTSANTOBAGCKKIFGSLA 375
 334 334 QY ----- 333

ATP (BY SIMILARITY).

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

GenCore version 5.1.6
Copyright (c) 1993, - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:29:10 ; Search time 18.0048 Seconds
(without alignments)
3013.200 Million cell updates/sec

Title: US-09-821-883-3
Perfect score: 3062
Sequence: 1 MRAAPLLARAASLSIGFLP.....ENPEYLGLDVPAAAHHHHH 564
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78: *
1: piri: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2417	78.9	1255	A24571	protein-tyrosine k
2	1962	64.1	1260	1 TVRTNU	p-185 precursor -
3	1946.5	63.6	1254	2 I48161	epidermal growth f
4	752	24.6	1223	1 TVCHLV	epidermal growth f
5	751	24.5	527	2 A42032	epidermal growth f
6	723.5	23.6	1308	2 A47253	epidermal growth f
7	719.5	23.5	1210	1 GQHUE	epidermal growth f
8	717.5	23.4	1210	2 A53183	epidermal growth f
9	713.5	23.3	644	2 A36325	kinase-related tra
10	697	22.8	1342	2 A36223	epidermal growth f
11	662	21.6	1339	2 JC4387	protein-tyrosine k
12	651	21.3	1166	1 S06142	epidermal growth f
13	539.5	17.6	843	2 A27131	epidermal growth f
14	404.5	13.2	1323	2 E88257	protein let-23 (im
15	404.5	13.2	1374	2 S70712	protein-tyrosine k
16	390.5	12.8	1369	2 S70713	protein-tyrosine k
17	379	12.4	1330	1 GQFFE	epidermal growth f
18	341	11.1	366	2 D45558	epidermal growth f
19	341	11.1	1717	1 A45558	epidermal growth f
20	331	10.8	333	2 B45558	epidermal growth f
21	331	10.8	342	2 C45558	epidermal growth f
22	318	10.4	1363	2 T43220	insulin-like growt
23	276.5	9.0	1477	2 T18534	protein-tyrosine k
24	271	8.9	2101	2 S57245	insulin receptor (
25	271	8.9	2148	1 A56081	insulin receptor -
26	258	8.4	1382	1 INHUR	insulin receptor p
27	252.5	8.2	1372	2 A34157	insulin-like growt
28	252	8.2	1607	2 T43212	insulin receptor p
29	249	8.1	1383	2 A36080	insulin receptor p

30	246.5	8.1	1300	2	A36502	insulin receptor-r
31	238.5	7.8	1367	1	IGHUR1	insulin-like growt
32	236.5	7.7	1371	2	A33837	insulin-like growt
33	235.5	7.7	183	2	JH0803	tyrosine kinase re
34	233.5	7.6	540	2	B47417	insulin receptor-r
35	231.5	7.6	1268	2	B36502	insulin receptor-r
36	224.5	7.3	329	2	A48805	insulin-like growt
37	207	6.8	1390	2	T30346	insulin receptor -
38	188	6.1	1846	2	J42047	insulin receptor h
39	168	5.5	386	1	JH0610	acid phosphatase (
40	155	5.1	5262	2	T03454	ALR protein - huma
41	153.5	5.0	707	2	A46302	PTB-associated epl
42	146	4.8	351	2	S50754	hypothetical prote
43	146	4.8	442	2	S50062	cell wall glycopro
44	139.5	4.6	473	2	S50755	hypothetical prote
45	139.5	4.6	597	2	S72468	probable transcrip

ALIGNMENTS

RESULT 1

A24571
protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human
N:Alternate names: c-erb-B-2 protein precursor; kinase-related transforming protein erb
C:Species: Homo sapiens (man)
C:Date: 25-Oct-1987 #sequence revision 06-Dec-1996 #text change 11-Jun-1999
C:Accession: A24571; A25491; A44188; B44188; I59509; I57622
R:Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T.;
Nature 319, 230-234, 1986
A:Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth i
A:Reference number: A24571; MUID:86118663; PMID:3003577
A:Accession: A24571
A:Molecule type: mRNA
A:Residues: 1-1255 <YAM>
A:Cross-references: GB:X03363; NID:g31197; PIDN:CAA27060.1; PID:g31198
R:Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985
A:Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epide
A:Reference number: A25491; MUID:86016729; PMID:2995967
A:Accession: A25491
A:Molecule type: DNA
A:Residues: 737-1031 <SEM>
A:Cross-references: GB:M11767; NID:g182163; PIDN:AAA35808.1; PID:g553282
R:Coussens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg,
Science 230, 1132-1139, 1985
A:Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chromo-
A:Reference number: A44188; MUID:86070181; PMID:2999974
A:Accession: A44188
A:Molecule type: DNA
A:Residues: 740-910 <COU1>
A:Cross-references: GB:M12036; NID:g183988; PIDN:AAA35978.1; PID:g183989
A:Accession: B44188
A:Molecule type: mRNA
A:Residues: 1-517, 'RALL', 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>
A:Cross-references: GB:M11730; NID:g183986
R:King, C.R.; Kraus, M.H.; Aaronson, S.A.
Science 229, 974-976, 1985
A:Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.
A:Reference number: I59509; MUID:85272597; PMID:2992089
A:Accession: I59509
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 832-909 <REX>
A:Cross-references: GB:L29395; NID:g459807; PIDN:AAA35809.1; PID:g459808
R:Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Glivol, D.
Mol. Cell. Biol. 7, 2597-2601, 1987
A:Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptional
A:Reference number: I57622; MUID:87286898; PMID:3039351
A:Accession: I57622
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-191 <TAL>

A:Cross-references: GB:M16792; NID:q183983; PIDN:AAA58637.1; PID:g553332
C:Comment: Amplification and overexpression of this erbB-related gene occurs in about 30
C:Genetics:
A:Gene: GDB:ERBB2; NGL; NEU; HER-2
A:Cross-references: GDB:120613; OMIM:164870
A:Map position: 17q21.1-17q21.1
A:Introns: 25/1; 75/3; 147/1; 883/3
A:Note: the list of introns is incomplete
C:Function:
A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
inase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-1255/Product: protein-tyrosine kinase erbB2 #status predicted <MAT>
F:22-653/Domain: extracellular #status predicted <EXT>
F:70-304/Domain: EGF receptor extracellular domain repeat <EE1>
F:395-605/Domain: EGF receptor extracellular domain repeat <EE2>
F:654-675/Domain: transmembrane #status predicted <TM>
F:676-1255/Domain: intracellular #status predicted <INT>
F:718-983/Domain: intracellular kinase homology <KIN>
F:726-734/Region: protein kinase ATP-binding motif
F:68.124.187,259,530,571,629/Binding site: carboxylate (Asn) (covalent) #status predicted
F:686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F:753/Active site: Lys #status predicted
F:1139,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 78.9%; Score 2417; DB 1; Length 1255;
Best Local Similarity 41.2%; Pred. No. 2.5e-135;
Matches 511; Conservative 4; Mismatches 6; Indels 718; Gaps 2;

QY 35 LARGASTQVCTGTMKRLRASPETHLDMLRHLYQGVQVQGNLELYLPTNASLSFLQ 94
DB 16 LPPGAASTQVCTGTMKRLRASPETHLDMLRHLYQGVQVQGNLELYLPTNASLSFLQ 75
QY 95 DIQEVQGVYLIAHNVROVPLQRLRIVRGTLQFEDNYALAVLDNGDPLNNTTPVTGASPG 154
DB 76 DIQEVQGVYLIAHNVROVPLQRLRIVRGTLQFEDNYALAVLDNGDPLNNTTPVTGASPG 135
QY 155 GLRELQLSLTILKGGVLIQRPOLCYQDTILWKIDIFHNKQLALTIDITNRSRACHPC 214
DB 136 GLRELQLSLTILKGGVLIQRPOLCYQDTILWKIDIFHNKQLALTIDITNRSRACHPC 195
QY 215 SPMCKGSRGWSSSDCSLTITVCAGGCARCKGLPTDCHEQCAAGCTGPKHSDCLAC 274
DB 196 SPMCKGSRGWSSSDCSLTITVCAGGCARCKGLPTDCHEQCAAGCTGPKHSDCLAC 255
QY 275 LHFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYLSLTDVGSASII- 333
DB 256 LHFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYLSLTDVGSCTLVC 315
QY 334 ----- 333
DB 316 PLHNQEVTAEDGTQCEKSKPCARVCYGLGMEHLREVRVTSANIQBPAGCKIFGSLA 375
QY 334 ----- 333
DB 376 FLFPESFDGPASNTAPLQEQVPELLEITGYLIYSAMPDLSLPSVQNLQVGRRI 435
QY 334 ----- 333
DB 436 LHNGAYSLTQLGIGISWLGRLSRLGSLALIHNTLHLCFVHTVPWDLFPNPHQALLH 495
QY 334 ----- 333
DB 496 TANRPEDECVGBGLACHQLCARGHCWPGFTQVCNCSQFLRQECVCECRVLQGLPREYV 555
QY 334 ----- 333
DB 556 NARHCLPCHPECPQNGSVTCFGBADQCACAHYKDPFCVARCPGKVDLSYMPIWK 615
QY 334 ----- 333

RESULT 2

TVRTWU

protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 11-Jun-1999

C:Accession: A24562; A61204

R:Batgmann, C.I.; Hung, M.C.; Weinberg, R.A.

Nature 319, 226-230, 1986

A:Title: The neu oncogene encodes an epidermal growth factor receptor-related protein.

A:Reference number: A24562; MUID:86118662; PMID:3945311

A:Accession: A24562

A:Molecule type: mRNA

A:Residues: 1-1260 <BAR>

A:Cross-references: EMBL:X03362; NID:g56745; PIDN:CAA27059.1; PID:g56746

R:Masui, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohen,

Carcinogenesis 12, 1975-1978, 1991

A:Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals no

2-thiazolyl]formamide or N-methyl-N-nitrosourea.

A:Reference number: A61204; MUID:92035293; PMID:1682063

A:Accession: A61204

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 637-663, 'Y', 665-702 <MAS>

A:Note: authors translated the codon GCA for residue 25 as Val

C:Genetics:

A:Gene: neu

C:Superfamily: epidermal growth factor receptor; protein kinase homology

C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:41:37 ; Search time 59.6409 Seconds
(without alignments)
3032.635 Million cell updates/sec

Title: US-09-821-883-3
Perfect score: 3062
Sequence: 1 MRAAPLLARASLSGLFLF.....ENPEYGLDVPAAHHHHH 564

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/CT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/CTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3062	100.0	564	US-09-821-883-3	Sequence 3, Appli
2	3014	98.4	697	US-09-821-883-4	Sequence 4, Appli
3	3005.5	98.2	555	US-09-821-883-1	Sequence 1, Appli
4	2957.5	96.6	690	US-09-821-883-2	Sequence 2, Appli
5	2585	84.4	919	US-09-854-356-6	Sequence 6, Appli
6	2422	79.1	1255	US-09-811-123-9	Sequence 9, Appli
7	2422	79.1	1255	US-09-811-115-3	Sequence 3, Appli
8	2422	79.1	1255	US-09-769-508-2	Sequence 2, Appli
9	2422	79.1	1255	US-09-984-092-4	Sequence 4, Appli
10	2422	79.1	1255	US-10-177-293-126	Sequence 126, App
11	2422	79.1	1255	US-10-207-498-6	Sequence 6, Appli
12	2422	79.1	1255	US-10-338-730-2	Sequence 2, Appli
13	2422	79.1	1255	US-10-323-892-4	Sequence 4, Appli
14	2422	79.1	1255	US-10-272-437A-28	Sequence 28, Appli
15	2422	79.1	1255	US-10-117-937-594	Sequence 594, App

16	2422	79.1	1255	15	US-10-435-696-36	Sequence 36, Appli
17	2422	79.1	1255	16	US-10-441-779C-4	Sequence 4, Appli
18	2422	79.1	1255	16	US-10-734-564-126	Sequence 126, App
19	2417	78.9	1255	9	US-09-854-356-1	Sequence 1, Appli
20	2417	78.9	1255	9	US-09-930-125-2	Sequence 2, Appli
21	2417	78.9	1255	10	US-09-441-411-6	Sequence 6, Appli
22	2417	78.9	1255	12	US-10-469-162-3	Sequence 3, Appli
23	2417	78.9	1255	12	US-10-253-286-553	Sequence 553, App
24	2417	78.9	1255	12	US-09-765-973-2	Sequence 2, Appli
25	2417	78.9	1255	12	US-10-418-027-3	Sequence 3, Appli
26	2417	78.9	1255	14	US-10-207-655-45	Sequence 45, Appli
27	2417	78.9	1255	14	US-10-313-644-2	Sequence 2, Appli
28	2417	78.9	1255	15	US-10-394-322A-17	Sequence 17, Appli
29	2417	78.9	1255	15	US-10-245-871-553	Sequence 553, App
30	2417	78.9	1255	15	US-10-149-138-4641	Sequence 4641, App
31	2417	78.9	1255	16	US-10-647-005-68	Sequence 68, Appli
32	2417	78.9	1255	16	US-10-149-138-4641	Sequence 4641, App
33	2410	78.7	1253	14	US-10-146-473-72	Sequence 72, Appli
34	1974.5	64.5	1256	9	US-09-854-356-14	Sequence 14, Appli
35	1962	64.1	1260	9	US-09-870-759-118	Sequence 118, App
36	1962	64.1	1260	10	US-09-751-708A-118	Sequence 118, App
37	1959.5	64.0	1256	9	US-09-854-356-2	Sequence 2, Appli
38	1838.5	60.0	479	9	US-09-821-883-5	Sequence 5, Appli
39	1614.5	52.7	645	9	US-09-921-161-1	Sequence 1, Appli
40	1614.5	52.7	645	14	US-10-268-501-13	Sequence 13, Appli
41	1614.5	52.7	645	15	US-10-608-626-13	Sequence 13, Appli
42	1614.5	52.7	653	9	US-09-854-356-3	Sequence 3, Appli
43	1614.5	52.7	685	15	US-10-412-804A-4	Sequence 4, Appli
44	1614.5	52.7	690	15	US-10-412-804A-11	Sequence 11, Appli
45	1614.5	52.7	712	9	US-09-854-356-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-821-883-3
; Sequence 3, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Gradiš, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; TITLE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500* construct
US-09-821-883-3

Query Match 100.0%; Score 3062; DB 9; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.3e-206;
Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MRAAPLLARASLSGLFLFLFFWLDRLAKELARGAASVCTGTGDKMLRPASPT 60	
Db	1	MRAAPLLARASLSGLFLFLFFWLDRLAKELARGAASVCTGTGDKMLRPASPT 60	
Qy	61	HLDMRLHYQGGVQVQGNLELYPTNASLSFLQDIQVQGVYLAHNRVQVPLQRLRI 120	
Db	61	HLDMRLHYQGGVQVQGNLELYPTNASLSFLQDIQVQGVYLAHNRVQVPLQRLRI 120	
Qy	121	VRGTQLFEDNVALVDNGDPLNNTTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOL 180	

Db 121 VRGTLQFEDNYALVLDNGDPLNNTTPTVTGASPGGLRELQLRSLTEILKGGVLIQRNPOL 180
QY 181 CYQDTILWKDIFHKNQNALTLIDTNRSRACHPCSPMCKGSRGWGSESSDCQSLTRTVCA 240
Db 181 CYQDTILWKDIFHKNQNALTLIDTNRSRACHPCSPMCKGSRGWGSESSDCQSLTRTVCA 240
QY 241 GGCARCKGFLPTDCCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFES 300
Db 241 GGCARCKGFLPTDCCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFES 300
QY 301 MNPEGRYTFGASCVTACPYNYLSTDVGSASIIINFEKLGAGGMVHRRSSSTRSGGDL 360
Db 301 MNPEGRYTFGASCVTACPYNYLSTDVGSASIIINFEKLGAGGMVHRRSSSTRSGGDL 360
QY 361 TLGLEPSEEEAPRSLAPSEAGSDVFDGDLGMGAAGLQSLPTHTDPSLQRYSDPTVP 420
Db 361 TLGLEPSEEEAPRSLAPSEAGSDVFDGDLGMGAAGLQSLPTHTDPSLQRYSDPTVP 420
QY 421 LPSETDGVVAPLTCSPQPEYVNPQVRPQPPSPREGPLPAARPAGATLERAKTILSPGKNG 480
Db 421 LPSETDGVVAPLTCSPQPEYVNPQVRPQPPSPREGPLPAARPAGATLERAKTILSPGKNG 480
QY 481 VVKOVFAFGGAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLVYWDQDPPERGAPSTPKG 540
Db 481 VVKOVFAFGGAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLVYWDQDPPERGAPSTPKG 540
QY 541 TPTAENPEYLGIDVPAHHHHH 564
Db 541 TPTAENPEYLGIDVPAHHHHH 564

RESULT 2

US-09-821-883-4

; Sequence 4, Application US/09821883

; Patent No. US20020061310A1

; GENERAL INFORMATION:

; APPLICANT: Laus, Reiner

; APPLICANT: Vidovic, Damir

; APPLICANT: Graddis, Thomas

; TITLE OF INVENTION: Compositions and Methods for Dendritic

; FILE REFERENCE: 7636-0022.30

; CURRENT APPLICATION NUMBER: US/09/821,883

; CURRENT FILING DATE: 2001-03-30

; PRIOR APPLICATION NUMBER: US 60/193,504

; PRIOR FILING DATE: 2000-03-30

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 697

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: HER500*-rCM-CSF construct

US-09-821-883-4

Query Match 98.4%; Score 3014; DB 9; Length 697;
Best Local Similarity 100.0%; Pred. No. 4.1e-203;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRAAPLLIARAASLSGLFLLFFWLDRLSVLAKELARGAASQTCTGDMKRLPASPET 60
Db 1 MRAAPLLIARAASLSGLFLLFFWLDRLSVLAKELARGAASQTCTGDMKRLPASPET 60
QY 61 HLDMLRHLYQCCVQVQGNLELTPTNASLSFLQDIQEVQGVLIHNRQVPLQRLRI 120
Db 61 HLDMLRHLYQCCVQVQGNLELTPTNASLSFLQDIQEVQGVLIHNRQVPLQRLRI 120
QY 121 VRGTLQFEDNYALVLDNGDPLNNTTPTVTGASPGGLRELQLRSLTEILKGGVLIQRNPOL 180
Db 121 VRGTLQFEDNYALVLDNGDPLNNTTPTVTGASPGGLRELQLRSLTEILKGGVLIQRNPOL 180

QY 181 CYQDTILWKDIFHKNQNALTLIDTNRSRACHPCSPMCKGSRGWGSESSDCQSLTRTVCA 240
Db 181 CYQDTILWKDIFHKNQNALTLIDTNRSRACHPCSPMCKGSRGWGSESSDCQSLTRTVCA 240
QY 241 GGCARCKGFLPTDCCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFES 300
Db 241 GGCARCKGFLPTDCCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFES 300
QY 301 MNPEGRYTFGASCVTACPYNYLSTDVGSASIIINFEKLGAGGMVHRRSSSTRSGGDL 360
Db 301 MNPEGRYTFGASCVTACPYNYLSTDVGSASIIINFEKLGAGGMVHRRSSSTRSGGDL 360
QY 361 TLGLEPSEEEAPRSLAPSEAGSDVFDGDLGMGAAGLQSLPTHTDPSLQRYSDPTVP 420
Db 361 TLGLEPSEEEAPRSLAPSEAGSDVFDGDLGMGAAGLQSLPTHTDPSLQRYSDPTVP 420
QY 421 LPSETDGVVAPLTCSPQPEYVNPQVRPQPPSPREGPLPAARPAGATLERAKTILSPGKNG 480
Db 421 LPSETDGVVAPLTCSPQPEYVNPQVRPQPPSPREGPLPAARPAGATLERAKTILSPGKNG 480
QY 481 VVKOVFAFGGAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLVYWDQDPPERGAPSTPKG 540
Db 481 VVKOVFAFGGAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLVYWDQDPPERGAPSTPKG 540
QY 541 TPTAENPEYLGIDVPAHHH 558
Db 541 TPTAENPEYLGIDVPAHHH 558

RESULT 3

US-09-821-883-1

; Sequence 1, Application US/09821883

; Patent No. US20020061310A1

; GENERAL INFORMATION:

; APPLICANT: Laus, Reiner

; APPLICANT: Vidovic, Damir

; APPLICANT: Graddis, Thomas

; TITLE OF INVENTION: Compositions and Methods for Dendritic

; FILE REFERENCE: 7636-0022.30

; CURRENT APPLICATION NUMBER: US/09/821,883

; CURRENT FILING DATE: 2001-03-30

; PRIOR APPLICATION NUMBER: US 60/193,504

; PRIOR FILING DATE: 2000-03-30

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 555

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: HER500 construct

US-09-821-883-1

Query Match 98.2%; Score 3005.5; DB 9; Length 555;
Best Local Similarity 98.4%; Pred. No. 1.2e-202;
Matches 555; Conservative 0; Mismatches 0; Indels 9; Gaps 1;
QY 1 MRAAPLLIARAASLSGLFLLFFWLDRLSVLAKELARGAASQTCTGDMKRLPASPET 60
Db 1 MRAAPLLIARAASLSGLFLLFFWLDRLSVLAKELARGAASQTCTGDMKRLPASPET 60
QY 61 HLDMLRHLYQCCVQVQGNLELTPTNASLSFLQDIQEVQGVLIHNRQVPLQRLRI 120
Db 61 HLDMLRHLYQCCVQVQGNLELTPTNASLSFLQDIQEVQGVLIHNRQVPLQRLRI 120
QY 121 VRGTLQFEDNYALVLDNGDPLNNTTPTVTGASPGGLRELQLRSLTEILKGGVLIQRNPOL 180
Db 121 VRGTLQFEDNYALVLDNGDPLNNTTPTVTGASPGGLRELQLRSLTEILKGGVLIQRNPOL 180
QY 181 CYQDTILWKDIFHKNQNALTLIDTNRSRACHPCSPMCKGSRGWGSESSDCQSLTRTVCA 240
Db 181 CYQDTILWKDIFHKNQNALTLIDTNRSRACHPCSPMCKGSRGWGSESSDCQSLTRTVCA 240

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:31:25 ; Search time 16.4294 Seconds
(without alignments)
1772.256 Million cell updates/sec

Title: US-09-821-883-3
Perfect score: 3062
Sequence: 1 MRAAPLLARAASLSGLFLF.....ENPEYGLDVPAAHHHHH 564

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2422	79.1	1255	4	US-09-527-487-2
2	2422	79.1	1255	4	US-09-811-115-3
3	2417	78.9	1255	1	US-08-467-083-68
4	2417	78.9	1255	1	US-08-414-4178-68
5	2417	78.9	1255	2	US-08-486-348A-68
6	2417	78.9	1255	2	US-08-625-101-2
7	2417	78.9	1255	2	US-08-468-545B-68
8	2417	78.9	1255	2	US-08-356-786-2
9	2417	78.9	1255	3	US-08-466-680B-68
10	2417	78.9	1255	4	US-09-354-533-68
11	2397	78.3	1255	2	US-08-484-438-8
12	1614.5	52.7	782	2	US-09-146-283-4
13	1614.5	52.7	782	3	US-08-579-823A-4
14	1614.5	52.6	419	4	US-09-344-195-4
15	1610.5	52.2	624	3	US-09-630-155-2
16	1599.5	52.2	624	3	US-08-422-108-1
17	1599.5	52.2	624	4	US-08-422-734-1
18	1177	38.4	580	1	US-08-414-4178-69
19	1177	38.4	580	2	US-08-486-348A-69
20	1177	38.4	580	2	US-08-468-545B-69
21	1177	38.4	580	3	US-08-466-680B-69
22	1177	38.4	580	4	US-09-354-533-69
23	860	28.1	166	4	US-09-648-067A-1
24	723.5	23.6	911	2	US-08-484-438-10
25	723.5	23.6	1058	2	US-08-484-438-4
26	723.5	23.6	1308	2	US-08-484-438-2
27	719.5	23.5	644	1	US-08-336-708A-9

28	719.5	23.5	1210	2	US-08-484-438-7	Sequence 7, Appli
29	719.5	23.5	1210	2	US-08-475-035-4	Sequence 4, Appli
30	701.5	22.9	478	4	US-09-570-454-2	Sequence 2, Appli
31	701.5	22.9	478	4	US-09-867-521-2	Sequence 2, Appli
32	697	22.8	1342	1	US-07-978-895-4	Sequence 4, Appli
33	697	22.8	1342	2	US-08-484-438-9	Sequence 9, Appli
34	697	22.8	1342	2	US-08-473-119-4	Sequence 4, Appli
35	697	22.8	1342	2	US-08-473-352-4	Sequence 4, Appli
36	697	22.8	1342	4	US-09-170-699-4	Sequence 4, Appli
37	697	22.8	1343	6	5183884-4	Patent No. 5183884
38	493	16.1	97	1	US-08-421-356-3	Sequence 3, Appli
39	493	16.1	97	4	US-09-046-783-3	Sequence 3, Appli
40	259	8.5	1382	2	US-08-737-715-2	Sequence 2, Appli
41	259	8.5	1382	2	US-09-457-040B-7	Sequence 7, Appli
42	238.5	7.8	516	3	US-08-746-559A-4	Sequence 4, Appli
43	238.5	7.8	1367	2	US-08-249-687C-2	Sequence 2, Appli
44	238.5	7.8	1367	2	US-08-625-819-2	Sequence 2, Appli
45	238.5	7.8	1367	3	US-08-746-559A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-527-487-2
; Sequence 2, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126881309200
; CURRENT APPLICATION NUMBER: US/09/527,487
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-527-487-2

Query Match	79.1%;	Score	2422;	DB	4;	Length	1255;
Best Local Similarity	41.3%;	Pred. No.	1.3e-18;				
Matches	512;	Conservative	4;	Mismatches	5;	Indels	718;
Gaps	2;						
Qy	35	LARGAASQVCTGTDMKRLRPASPTHLDMLRHLYQGCGVQVGNLELTLPNNSLSFIQ	94				
Db	16	LPPGAASQVCTGTDMKRLRPASPTHLDMLRHLYQGCGVQVGNLELTLPNNSLSFIQ	75				
Qy	95	DIQEVQGYVLIHAHQVQVPLQRLRIVRGTQLFEDNYALVLDNGDPLNNTTPTVTGASPG	154				
Db	76	DIQEVQGYVLIHAHQVQVPLQRLRIVRGTQLFEDNYALVLDNGDPLNNTTPTVTGASPG	135				
Qy	155	GLRELQLSLTELKGGVLIQNPOLCYQDTILWKDIFHKNNQLALTLDITNRSRACHPC	214				
Db	136	GLRELQLSLTELKGGVLIQNPOLCYQDTILWKDIFHKNNQLALTLDITNRSRACHPC	195				
Qy	215	SPMKGSRGWSGSSSDCSLRTVTCAGGCARCKGKPLPTDCHEQCAAGCTGPKHSCLAC	274				
Db	196	SPMKGSRGWSGSSSDCSLRTVTCAGGCARCKGKPLPTDCHEQCAAGCTGPKHSCLAC	255				
Qy	275	LHFNHSGICELHCPALVYNTDTTFESMPNPEGRYTFGASCVTACPNYLSLTDVGSASII-	333				
Db	256	LHFNHSGICELHCPALVYNTDTTFESMPNPEGRYTFGASCVTACPNYLSLTDVGSASII-	315				
Qy	334	-----	333				
Db	316	PLHNOEVTAEADGTQRCCKSPCARVCYGLGMEHLREVRVTSANIQEPAGCKKIFGSUA	375				
Qy	334	-----	333				
Db	376	FLPESFDGDPASNTAPLQEQVFETLEETGYLYISAWPDSLPLDSVFNQLVIRGRI	435				

```
QY 334 ----- 333
Db 436 LHNGAYSLTLOGLIGISWGLRSLRSLGSLALIHNNHNLCPVHTVPWDLFRNPHQALLH 495
QY 334 ----- 333
Db 496 TANRPEDECVGEGLAGHQLCARGHCWGPGPTQCVNCSQFLRQECVEBRCVLQGLPREVY 555
QY 334 ----- 333
Db 556 NARHCLPCHPECPQNGSVTCFGEADQCVACAHYKDPFCVACPSGVKPDLSYMPIWK 615
QY 334 ----- 333
Db 616 FPDEGACQPCINCTHSCVDLDDKGCAPAEORASPLTSIVSAVVGILLVVVLGVVFGILI 675
QY 334 ----- 333
Db 676 KRRQOKIRKYTMRRLLQETELVEPLTPSGAMPNQAQMRILKETELRKVKVLGSGAGFTVY 735
QY 334 ----- 333
Db 736 KGIWIPDGENVKIPVAIKVLRENTSPKANKEILDEAYVWAGVGSPPVSRLLIGLICLTSTVQ 795
QY 334 ----- 333
Db 796 LVTQLMPYGCILLDHYENRGRGLSODLLNWCQIAKMSYLEDVPLVHRDLAARNVLKVS 855
QY 334 ----- 333
Db 856 PNHVKITDFGLARLLDIDETEHADGKGKVPKIMWMALESILRRRFTHQSDVMSYGVTVWEL 915
QY 334 ----- 338
Db 916 MTFGAKPYDGIPAREIPDLLEKGERLPPOPICTIDVYIMVKWMIDSECRPRPRELVSE 975
QY 339 ----- 338
Db 976 FSRMARDPQRFVVIQNEIDLGPASPLDSTFYRSLLEDDMGDLVDAEYLVQQGFFCPDP 1035
QY 339 --GAGGMVHRRHSSTSGGDLTLGLEPSEEAAPRSLAPSEGAGSDVFDGLGMGAA 396
Db 1036 AFGAGMVHRRHSSTSGGDLTLGLEPSEEAAPRSLAPSEGAGSDVFDGLGMGAA 1095
QY 397 KGLQSLPHDPSLQRYSEDPTVLPSETDGVVAPLTCSPQPEYVQPDVBPQPSREG 456
Db 1096 KGLQSLPHDPSLQRYSEDPTVLPSETDGVVAPLTCSPQPEYVQPDVBPQPSREG 1155
QY 457 PLPAARPAATILERAKTILSPGKNGVYKDVFAFGGAVENPEYLTPQGGAAPOPHPPAFSP 516
Db 1156 PLPAARPAATILERAKTILSPGKNGVYKDVFAFGGAVENPEYLTPQGGAAPOPHPPAFSP 1215
QY 517 AFDNLYYWDQPPPERGAPSTFKGPTTAENPEYLGLDVP 555
Db 1216 AFDNLYYWDQPPPERGAPSTFKGPTTAENPEYLGLDVP 1254
```

RESULT 2

```
US-09-811-115-3
; Sequence 3, Application US/09811115
; Patent No. 6632979
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwalli, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT. 034A
; CURRENT APPLICATION NUMBER: US/09/811,115
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 3
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-115-3

Query Match 79.1%; Score 2422; DB 4; Length 1255;
Best Local Similarity 41.3%; Pred. No. 1.3e-187;
Matches 512; Conservative 4; Mismatches 5; Indels 718; Gaps 2;

QY 35 LARGAASQVCTGDMKRLRLPASBETHLDMRLPHLYQGCQVQGNLELYLPTNASLSFLQ 94
Db 16 LPPGAASQVCTGDMKRLRLPASBETHLDMRLPHLYQGCQVQGNLELYLPTNASLSFLQ 75
QY 95 DIOEQVQGVLLTAHNOVRQVPLQRLRIVRGQTOLFEEDNYALAVLDNGDPLNNTTPTVTGASPG 154
Db 76 DIOEQVQGVLLTAHNOVRQVPLQRLRIVRGQTOLFEEDNYALAVLDNGDPLNNTTPTVTGASPG 135
QY 155 GLRELQRLSLTEILKGGVLIQRNPLQCYQDITLWKDIFHKNNQLALTLIDTNRSRACHPC 214
Db 136 GLRELQRLSLTEILKGGVLIQRNPLQCYQDITLWKDIFHKNNQLALTLIDTNRSRACHPC 195
QY 215 SPMCKGSRGWESSEDCQSLTRTVACGACRCKGPLPTDCCHEQCAAGCTGPKHSDCLAC 274
Db 196 SPMCKGSRGWESSEDCQSLTRTVACGACRCKGPLPTDCCHEQCAAGCTGPKHSDCLAC 255
QY 275 LHFNHSGICELHCPALVTYNTDTPESMPNPEGRYTFGASCVTACPYNYLSTDVGSASII- 333
Db 256 LHFNHSGICELHCPALVTYNTDTPESMPNPEGRYTFGASCVTACPYNYLSTDVGSCTLVC 315
QY 334 ----- 333
Db 316 PLHNQEVTAEDGTQRCCKSKPCARVCYGLGMEHLREVRAVTSANIQEFAGCKKIFGSLA 375
QY 334 ----- 333
Db 376 FLPSFDDGPASNTAPLQPEQLQVFETLEBITGYLIYISAWPDSLPLDSVFQNLQVIRGI 435
QY 334 ----- 333
Db 436 LHNGAYSLTLOGLIGISWGLRSLRSLGSLALIHNNHNLCPVHTVPWDLFRNPHQALLH 495
QY 334 ----- 333
Db 496 TANRPEDECVGEGLAGHQLCARGHCWGPGPTQCVNCSQFLRQECVEBRCVLQGLPREVY 555
QY 334 ----- 333
Db 556 NARHCLPCHPECPQNGSVTCFGEADQCVACAHYKDPFCVACPSGVKPDLSYMPIWK 615
QY 334 ----- 333
Db 616 FPDEGACQPCINCTHSCVDLDDKGCAPAEORASPLTSIVSAVVGILLVVVLGVVFGILI 675
QY 334 ----- 333
Db 676 KRRQOKIRKYTMRRLLQETELVEPLTPSGAMPNQAQMRILKETELRKVKVLGSGAGFTVY 735
QY 334 ----- 333
Db 736 KGIWIPDGENVKIPVAIKVLRENTSPKANKEILDEAYVWAGVGSPPVSRLLIGLICLTSTVQ 795
QY 334 ----- 333
Db 796 LVTQLMPYGCILLDHYENRGRGLSODLLNWCQIAKMSYLEDVPLVHRDLAARNVLKVS 855
QY 334 ----- 333
Db 856 PNHVKITDFGLARLLDIDETEHADGKGKVPKIMWMALESILRRRFTHQSDVMSYGVTVWEL 915
QY 334 ----- 338
Db 916 MTFGAKPYDGIPAREIPDLLEKGERLPPOPICTIDVYIMVKWMIDSECRPRPRELVSE 975
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:20:15 ; Search time 62.1165 Seconds
(without alignments)
2565.449 Million cell updates/sec

Title: US-09-821-883-3
Perfect score: 3062
Sequence: 1 MRAAPLLARASLSGLF.....ENPEYGLDVPAAAHHHHH 564

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107.

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003s.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3062	100.0	564	AAE13110	Human HER
2	3014	98.4	697	AAE13111	Human HER
3	3005.5	98.2	555	AAE13108	Human HER
4	2957.5	96.6	690	AAE13109	Human HER
5	2585	84.4	919	AAE13103	Human HER
6	2585	84.4	919	AAE13103	Human HER
7	2422	79.1	1255	AAE13108	Human HER
8	2422	79.1	1255	AAE13108	Human HER
9	2422	79.1	1255	AAE13108	Human HER
10	2422	79.1	1255	AAE13108	Human HER
11	2422	79.1	1255	AAE13108	Human HER
12	2422	79.1	1255	AAE13108	Human HER
13	2422	79.1	1255	AAE13108	Human HER
14	2422	79.1	1255	AAE13108	Human HER
15	2422	79.1	1255	AAE13108	Human HER
16	2422	79.1	1255	AAE13108	Human HER
17	2422	79.1	1255	AAE13108	Human HER
18	2422	79.1	1255	AAE13108	Human HER
19	2417	78.9	1255	AAE13108	Human HER
20	2417	78.9	1255	AAE13108	Human HER
21	2417	78.9	1255	AAE13108	Human HER
22	2417	78.9	1255	AAE13108	Human HER
23	2417	78.9	1255	AAE13108	Human HER
24	2417	78.9	1255	AAE13108	Human HER
25	2417	78.9	1255	AAE13108	Human HER

26	2417	78.9	1255	5	AAU77114	Human Her
27	2417	78.9	1255	5	AAU77114	Human Her
28	2417	78.9	1255	5	AAU77114	Human Her
29	2417	78.9	1255	6	ABR43687	Human C-e
30	2417	78.9	1255	6	ABR43687	Human Her
31	2417	78.9	1255	7	ADC09593	Her2/Neu
32	2417	78.9	1255	7	ADC09593	Binding d
33	2417	78.9	1255	7	ADC09593	Human Pro
34	2417	78.9	1255	7	ADC09593	Human Her
35	2410	78.7	1253	7	ADC35106	Human bre
36	2401	78.4	1433	2	AAU98923	Sequence
37	2346	76.6	1223	5	AAU98923	Human bre
38	2142.5	70.0	920	5	AAU98923	Mouse Her
39	2142.5	70.0	920	5	AAU98923	Mouse Her
40	2104	68.7	1200	3	AAE13110	Human HER
41	1974.5	64.5	1256	3	AAE13110	Human HER
42	1974.5	64.5	1256	3	AAE13110	Human HER
43	1974.5	64.5	1256	3	AAE13110	Human HER
44	1962	64.1	1260	6	AAU79135	HER-2 (HE
45	1959.5	64.0	1256	3	AAE13110	Rat HER-2

ALIGNMENTS

RESULT 1
AAE13110
ID AAE13110 standard; protein; 564 AA.
XX
AC AAE13110;
XX
DT 28-JAN-2002 (first entry)
XX
DE Human HER500 fusion protein construct comprising OVA-derived octapeptide.
XX
KW Immunostimulatory fusion protein; IFP; antigen component; therapy;
KW Immunostimulatory component; T-cell mediated immune response; DC;
KW Dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human; OVA;
KW HER-2 protein; ovalbumin-derived octapeptide; HER500 fusion protein.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
PN WO200174855-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US010515.
XX
PR 30-MAR-2000; 2000US-0193504P.
XX
PA (DEND-) DENDREON CORP.
XX
PI Laus R, Vidovic D, Graddis T;
XX
WPI: 2001-662965/76.
XX
N-PSDB; AAD21566.

An immunostimulatory fusion protein comprising the intracellular domain of HER-2 and an antigen elicits an immune response to the antigen and is useful for the treatment of associated cancer associated.

Claim 7; Page 26; 59pp; English.

The invention relates to immunostimulatory fusion proteins (IFP) and nucleic acid molecules encoding such proteins. The IFPs comprise a polypeptide antigen component and an immunostimulatory domain derived from the intracellular domain of HER-2 protein which is effective to elicit a protective dendritic cell (DC)-induced T-cell mediated cellular immune response to the antigen. IFP or superactivated dendritic cells are

used to treat cancer e.g. breast carcinoma, ovarian and colon cancer associated with a particularly antigen. The present sequence is HER500 fusion protein construct which comprises human PAP signal sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal sequence, mature HER-2 membrane distal extracellular domain, an Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide, HER-2 membrane distal intracellular domain and a C-terminal tag

Sequence 564 AA;

Query Match 100.0%; Score 3062; DB 4; Length 564;
 Best Local Similarity 100.0%; Pred. No. 1.8e-211;
 Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAAPLLARAAASLSLGLFLLFFWLDRLSVLAKELARGAASQVCTGTDMLRLPASPET 60
 DB 1 MRAAPLLARAAASLSLGLFLLFFWLDRLSVLAKELARGAASQVCTGTDMLRLPASPET 60

QY 61 HLDMLRHLVYQCCVQVQGNLELTYLPTNASLSFLQDIQEVQGVLIHNRVQVPLQLRLI 120
 DB 61 HLDMLRHLVYQCCVQVQGNLELTYLPTNASLSFLQDIQEVQGVLIHNRVQVPLQLRLI 120

QY 121 VRGTQLEFDNYALVDNGDPLNNTPTVTGASPGGLRELQLRSLTEILKGGVLIQRPQL 180
 DB 121 VRGTQLEFDNYALVDNGDPLNNTPTVTGASPGGLRELQLRSLTEILKGGVLIQRPQL 180

QY 181 CYQDTILWKDIFHKNQNLALTIDNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVCA 240
 DB 181 CYQDTILWKDIFHKNQNLALTIDNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVCA 240

QY 241 GGCARCKGPLEPTDCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFES 300
 DB 241 GGCARCKGPLEPTDCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFES 300

QY 301 MPNPEGRYTFGASCVTACPYNYLSTDVGSASIIINFEKLGAGGMVHRRSSSTRSGGDL 360
 DB 301 MPNPEGRYTFGASCVTACPYNYLSTDVGSASIIINFEKLGAGGMVHRRSSSTRSGGDL 360

QY 361 TLGLEPSEEEAPRSLAPSEAGSDVFDGLGMGAAGKLSLPTHDPSPLQRYSDPTVP 420
 DB 361 TLGLEPSEEEAPRSLAPSEAGSDVFDGLGMGAAGKLSLPTHDPSPLQRYSDPTVP 420

QY 421 LPSETDGYVAPLTCSPQEVYVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 480
 DB 421 LPSETDGYVAPLTCSPQEVYVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 480

QY 481 VVKQVFAFGGAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLVYWDQDPPERGAPESTPKG 540
 DB 481 VVKQVFAFGGAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLVYWDQDPPERGAPESTPKG 540

QY 541 TPTAENPEYLGIDVPAHAHHHHH 564
 DB 541 TPTAENPEYLGIDVPAHAHHHHH 564

RESULT 2

AAE13111
 ID AAE13111 standard; protein; 697 AA.
 XX AAE13111;
 AC AAE13111;
 XX 28-JAN-2002 (first entry)
 XX Human HER500-rGM-CSF fusion construct comprising OVA-derived peptide.
 DE Immunostimulatory fusion protein; IFP; antigen component; therapy;
 XX immunostimulatory component; T-cell mediated immune response; DC;
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
 KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
 KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
 KW HER-2 protein; granulocyte-macrophage colony stimulating factor;
 KW ovalbumin-derived octapeptide; OVA; rat; HER500-rGM-CSF fusion protein.

OS Homo sapiens.
 OS Rattus norvegicus.
 OS Unidentified.
 OS Chimeric.
 XX WO200174855-A2.
 PN 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US010515.
 PF 30-MAR-2000; 2000US-0193504P.
 PR (DEND-) DENDREON CORP.
 PA Laus R. Vidovic D. Graddis T;
 PI WPI; 2001-662965/76.
 XX N-PSDB; AAD21567.
 DR An immunostimulatory fusion protein comprising the intracellular domain of HER-2 and an antigen elicits an immune response to the antigen and is useful for the treatment of associated cancer associated.

PS Claim 7; Page 27; 59pp; English.

The invention relates to immunostimulatory fusion proteins (IFP) and nucleic acid molecules encoding such proteins. The IFPs comprise a polypeptide antigen component and an immunostimulatory component derived from the intracellular domain of HER-2 protein which is effective to elicit a protective dendritic cell (DC)-induced T-cell mediated cellular immune response to the antigen. IFP or superactivated dendritic cells are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer associated with a particularly antigen. The present sequence is HER500 rGM-CSF fusion protein construct which comprises human PAP signal sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal sequence, mature HER-2 membrane distal extracellular domain, an Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide, HER-2 membrane distal intracellular domain, an Ala linker, a mature rat granulocyte-macrophage colony stimulating factor (GM-CSF) sequence and a C-terminal tag

Sequence 697 AA;

Query Match 98.4%; Score 3014; DB 4; Length 697;
 Best Local Similarity 100.0%; Pred. No. 6.7e-208;
 Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAAPLLARAAASLSLGLFLLFFWLDRLSVLAKELARGAASQVCTGTDMLRLPASPET 60
 DB 1 MRAAPLLARAAASLSLGLFLLFFWLDRLSVLAKELARGAASQVCTGTDMLRLPASPET 60

QY 61 HLDMLRHLVYQCCVQVQGNLELTYLPTNASLSFLQDIQEVQGVLIHNRVQVPLQLRLI 120
 DB 61 HLDMLRHLVYQCCVQVQGNLELTYLPTNASLSFLQDIQEVQGVLIHNRVQVPLQLRLI 120

QY 121 VRGTQLEFDNYALVDNGDPLNNTPTVTGASPGGLRELQLRSLTEILKGGVLIQRPQL 180
 DB 121 VRGTQLEFDNYALVDNGDPLNNTPTVTGASPGGLRELQLRSLTEILKGGVLIQRPQL 180

QY 181 CYQDTILWKDIFHKNQNLALTIDNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVCA 240
 DB 181 CYQDTILWKDIFHKNQNLALTIDNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVCA 240

QY 241 GGCARCKGPLEPTDCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFES 300
 DB 241 GGCARCKGPLEPTDCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFES 300

QY 301 MPNPEGRYTFGASCVTACPYNYLSTDVGSASIIINFEKLGAGGMVHRRSSSTRSGGDL 360
 DB 301 MPNPEGRYTFGASCVTACPYNYLSTDVGSASIIINFEKLGAGGMVHRRSSSTRSGGDL 360

QY 361 TLGLEPSEEEAPRSLAPSEAGSDVFDGLGMGAAGKLSLPTHDPSPLQRYSDPTVP 420

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:57:41 ; Search time 4652.68 Seconds
(without alignments)
4428.608 Million cell updates/sec

Title: US-09-821-883-2
Perfect score: 3739
Sequence: 1 MRAAPLLARAASLSGLF.....EPVQEGAPPPPAHHHHH 690
Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Command line parameters: -DEV=xlh
-Q=/sgn2_1/USPTO_spool/US09821883/runat_09092004_105127_6643/app_query.fasta_1.3100
-DB=EST -QFWI=fastap -SUFFIX=stc -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORESPT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09821883 @CGN 1 1 11571 @runat_09092004_105127_6643 -NCPU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT_DSPLCK=100 -LONGLOG
-DE_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_inv.*
18: em_gss_hum.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rtd.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1961	52.4	4323	11	AK031099	Mus muscu
2	1944.5	52.0	4463	11	AK083669	Mus muscu
3	1288	34.4	1016	12	BM562913	AGENCOURT
4	1176.5	31.5	1129	12	BM802792	AGENCOURT
5	1175.5	31.4	885	14	CA455074	AGENCOURT
6	1170	31.3	871	14	CA488274	AGENCOURT
7	1163	31.1	4715	11	AF318349	Homo sapi
8	1145.5	30.6	893	14	CA455141	AGENCOURT
9	1098.5	29.4	894	14	CA454131	AGENCOURT
10	1096	29.3	1004	14	CD515356	AGENCOURT
11	1087	29.1	653	12	BM721340	UI-E-E01-
12	1079	28.9	902	14	CA488868	AGENCOURT
13	1073	28.7	588	13	EX478931	DKP2p686A
14	1046	28.0	569	9	AL701765	DKF2p686P
15	1027	27.5	583	13	EX479114	DKF2p686P
16	1016.5	27.2	795	14	CA328613	UI-M-PY0-
17	1009.5	27.0	3110	11	AK031542	Mus muscu
18	1004	26.9	808	14	CA489534	AGENCOURT
19	997	26.7	616	14	CD721801	OJ01f07.Y
20	976	26.1	943	13	BQ958632	AGENCOURT
21	974	26.0	1988	11	BC023725	Mus muscu
22	974	26.0	3372	11	BC046553	Mus muscu
23	933	25.0	855	13	BU594980	AGENCOURT
24	924	24.7	813	12	BI557797	603236977
25	915.5	24.5	852	14	CD516283	AGENCOURT
26	915	24.5	614	10	AW370693	QVI-BT026
27	912	24.4	791	12	BI154872	602902857
28	903	24.2	717	14	CB598701	AGENCOURT
29	901	24.1	491	12	BM790293	K-EST0070
30	893	23.9	906	14	CA454570	AGENCOURT
31	889.5	23.8	998	12	BI649877	603296516
32	881	23.6	685	9	AU123871	AU123871
33	874.5	23.4	932	14	CA487981	AGENCOURT
34	868	23.2	609	10	AG701942	uq93b02.Y
35	864.5	23.1	621	12	BG283493	602407782
36	853	22.8	887	13	BQ717097	AGENCOURT
37	848	22.7	786	12	BI155788	602904360
38	835.5	22.3	649	9	AA496412	AA496412
39	817	21.9	533	14	CB437620	685491.MA
40	817	21.9	552	12	BG277542	ux45d07.Y
41	811	21.7	670	9	AI906012	RC-BT105-
42	808	21.6	876	13	BQ769889	UI-M-RT0-
43	803	21.5	887	14	CA980253	AGENCOURT
44	802	21.4	514	10	BF998814	QVO-GN014
45	792.5	21.2	678	12	BI555157	603236396

ALIGNMENTS

RESULT 1
AK031099
LOCUS
DEFINITION
Mus musculus 13 days embryo forelimb cDNA, RIKEN full-length enriched library, clone:530404N10 product:v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian), full insert sequence.
ACCESSION
AK031099
VERSION
AK031099.1 GI:26082143
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AK031099 4323 bp mRNA linear HTC 18-SEP-2003

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 15:10:16 ; Search time 834.278 Seconds
(without alignments)
4161.385 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739

Sequence: 1 MRAAPLLIARAASLSGLF.....EPVGEAGPPPPAAHHHHH 690

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
	Ygapop 10.0	Ygapext 0.5
	Fgapop 6.0	Fgapext 7.0
	Delop 6.0	Delext 7.0

Searched: 3304383 seqs, 2515761380 residues

Total number of hits satisfying chosen parameters: 6608766

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09821883/runat_09092004_105129_6736/app.query.fasta_1.3100
-DB=Published Applications NA -QPM=fastap -SUFFX=rnnpb -MINMATCH=0.1
-LOOFCU=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=DIOSUM62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09821883 @CGN 1 1 1949 @runat_09092004_105129_6736
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1:	/cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
2:	/cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq.*
3:	/cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq.*
4:	/cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq.*
5:	/cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq.*
6:	/cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq.*
7:	/cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq.*
8:	/cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
9:	/cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq.*
10:	/cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq.*
11:	/cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq.*
12:	/cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
13:	/cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
14:	/cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq.*
15:	/cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq.*
16:	/cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq.*
17:	/cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq.*
18:	/cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq.*
19:	/cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	ID	Description
------------	-------	--------------	----	-------------

ALIGNMENTS

RESULT 1

```
US-09-821-883-7
; Sequence 7, Application US/09821883
; Patent No. US2002061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821.883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 2070
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500-hGM-CSF construct
```

1	3739	100.0	2070	9	US-09-821-883-7	Sequence 7, Appl
2	3473.5	92.9	2091	9	US-09-821-883-9	Sequence 9, Appl
3	2974	79.5	1665	9	US-09-821-883-6	Sequence 6, Appl
4	2959.5	79.2	1692	9	US-09-821-883-8	Sequence 8, Appl
5	2426	64.9	4543	9	US-09-769-508-1	Sequence 1, Appl
6	2423	64.8	4530	9	US-09-877-177-11	Sequence 11, Appl
7	2423	64.8	4530	13	US-10-392-113-45	Sequence 45, Appl
8	2423	64.8	4530	13	US-10-426-836-11	Sequence 11, Appl
9	2423	64.8	4530	15	US-10-177-293-125	Sequence 125, App
10	2423	64.8	4530	15	US-10-007-926A-119	Sequence 119, App
11	2423	64.8	4530	15	US-10-338-730-1	Sequence 1, Appl
12	2423	64.8	4530	15	US-10-101-510-124	Sequence 124, App
13	2423	64.8	4530	16	US-10-116-275-131	Sequence 131, App
14	2423	64.8	4530	16	US-10-272-437A-27	Sequence 27, Appl
15	2423	64.8	4530	16	US-10-117-937-595	Sequence 595, App
16	2423	64.8	4530	16	US-10-159-563-208	Sequence 208, App
17	2423	64.8	4530	16	US-10-435-696-10	Sequence 10, Appl
18	2423	64.8	4530	17	US-10-734-564-59	Sequence 59, Appl
19	2423	64.8	4642	15	US-10-198-846-10896	Sequence 10896, A
20	2422	64.8	9274	9	US-09-811-123-7	Sequence 7, Appl
21	2422	64.8	9274	9	US-09-811-115-1	Sequence 1, Appl
22	2418	64.7	4473	10	US-09-441-411-5	Sequence 5, Appl
23	2418	64.7	4473	15	US-10-146-473-32	Sequence 32, Appl
24	2418	64.7	4473	15	US-10-207-655-44	Sequence 44, Appl
25	2418	64.7	4473	15	US-10-101-510-81	Sequence 81, Appl
26	2410.5	64.5	3765	15	US-10-207-498-5	Sequence 5, Appl
27	2410.5	64.5	3768	9	US-09-811-123-8	Sequence 8, Appl
28	2410.5	64.5	3768	9	US-09-811-115-2	Sequence 2, Appl
29	2410.5	64.5	3768	13	US-09-984-092-3	Sequence 3, Appl
30	2410.5	64.5	3768	13	US-10-280-576-3	Sequence 3, Appl
31	2410.5	64.5	3768	17	US-10-441-799C-3	Sequence 3, Appl
32	2405.5	64.3	3768	9	US-09-854-356-9	Sequence 9, Appl
33	2405.5	64.3	3768	9	US-09-930-125-1	Sequence 1, Appl
34	2405.5	64.3	3768	13	US-09-765-973-1	Sequence 1, Appl
35	2405.5	64.3	3768	15	US-10-313-644-1	Sequence 1, Appl
36	2393	64.0	4606	10	US-09-971-392-70	Sequence 70, Appl
37	2188.5	58.5	1437	9	US-09-821-883-10	Sequence 10, Appl
38	1968	52.6	3955	9	US-09-870-759-117	Sequence 117, App
39	1968	52.6	3955	9	US-09-854-356-10	Sequence 10, Appl
40	1968	52.6	3955	10	US-09-751-708A-117	Sequence 117, App
41	1961	52.4	3771	9	US-09-854-356-11	Sequence 11, Appl
42	1684.5	44.5	2132	16	US-10-412-804A-3	Sequence 3, Appl
43	1659	44.4	2164	16	US-10-412-804A-5	Sequence 5, Appl
44	1656	44.3	2149	16	US-10-412-804A-9	Sequence 9, Appl
45	1587	42.4	867	9	US-09-821-883-24	Sequence 24, Appl

US-09-821-883-7

Alignment Scores:

Pred. No.: 0 Length: 2070
Score: 3739.00 Matches: 690
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-821-883-2 (1-690) x US-09-821-883-7 (1-2070)

QY 1 MetArgAlaProLeuLeuAlaArgAlaLaserLeuLeuGlyPheLeuPhe 20
DB 1 ATGAGAGCTGACCCCTCTCTGCGCAGGAGCAGCAAGCCTTAGCCTTGGCTTTCTTGTTT 60
QY 21 LeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyValAla 40
DB 61 CTGCTTTTCTGGCTAGACCGAAGTGTACTAGCCAGGAGTTGGCGCGCGGCGCGCG 120
QY 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
DB 121 TCGACCCCAAGTGTGCACCGCACACATGAACTGGCGCTCCCTGGCCAGTCCCGAGACC 180
QY 61 HisLeuAspMetLeuArgHisLeuTyrglnGlyCysGlnValValGlnGlyAsnLeuGlu 80
DB 181 CACCTGGACATGCTCCCGCACCTCTACCGAGGCTGCCAGGTGGTGAGGGAACCTGGAA 240
QY 81 LeuThrTyrrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGlnValGln 100
DB 241 CTCACCTACCTGCCCAACCAATGCCAGCTGTCTCTCTGCCAGATATCCAGGAGTGCAG 300
QY 101 GlyTyrrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
DB 301 GGCTACGTGCTCATCGCTCACAACCAAGTAGGCGAGGTCCCACTGCAGAGCTCGGATT 360
QY 121 ValArgGlyThrGlnLeuPheGluAspAsnTyrrAlaLeuAlaValLeuAspAsnGlyAsp 140
DB 361 GTGGAGGACACCCAGCTCTTTGAGGACAACTATGCCCTTGGCCGTGTAGACAAATGGAGAC 420
QY 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln 160
DB 421 CCGCTGAACATACCAACCCCTGTACAGGGGCTCCCGCAGGAGGCTCGCGAGCTGCAG 480
QY 161 LeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeu 180
DB 481 CTTGGAAGCTTCACAGAGATCTGAAAGAGGGGTCTTGATCCAGCGGAACCCCGAGCTC 540
QY 181 CysTyrrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200
DB 541 TGCTACCGAGGACAGATTTTGTGAGAGACATCTTCCAAAGAACCAACAGCTGGCTCTC 600
QY 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220
DB 601 AACTGTAGACACCAACCGCTCTCGGGCTGCCACCCCTCTCTCCGATGTGTAAGGCG 660
QY 221 SerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAla 240
DB 661 TCCGCTGTCTGGGAGAGATTTCTGAGGATTTGTGAGAGCTGTGAGCGCTGTCTGTGCC 720
QY 241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCysAla 260
DB 721 GGTGGCTGTGCGCGCTGCAAGGGGCACTGCCCACTGTACTGTCCATGAGCAGTGTGCT 780
QY 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
DB 781 GCGGCTGCACGGGCCCCAAGCACTCTGACTGCTGCCCTGGCTCCCTCCACTTCAACCAAGT 840
QY 281 GlyIleCysGluLeuHisCysProAlaLeuValThrTyrrAsnThrAspThrPheGluSer 300
DB 841 GGCATCTGTGAGCTGCATGCGCCGCTGCTGCTACCTACAAACACAGACACGCTTTGAGTCC 900
QY 301 MetProAsnProGluGlyArgTyrrThrPheGlyAlaSerCysValThrAlaCysProTyrr 320

DB 901 ATGCCCAATCCGAGGGCGGTATACATTGGCGCCAGCTGTGTGACTGCTGTGCTCCCTAC 960
QY 321 AsnTyrrLeuSerThrAspValGlySerGlyAlaGlyGlyMetValHisHisArgHisArg 340
DB 961 AACTTACCTTTCTACCGACGTGGATCGGGCGCTGGGGCATGTGTCACACAGACACCGC 1020
QY 341 SerSerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGlu 360
DB 1021 AGCTCATCTACCGAGGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAG 1080
QY 361 GluAlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGly 380
DB 1081 GAGGCCCCAGGCTCTCCTACCTGGCACCTCCGAGGGGCTGGCTCCGATGTATTGATGGT 1140
QY 381 AspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerPro 400
DB 1141 GACCTGGGAATGGGGGAGCAAGGGCTGCAAGGCTCCCAACACACATGACCCCGAGCCT 1200
QY 401 LeuGlnArgTyrrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrrVal 420
DB 1201 CTACGCGGTACAGTGAGGACCCCACTACCTCCCTCTGAGACTGTGCTGCTAGTT 1260
QY 421 AlaProLeuThrCysSerProGlnProGluTyrrValAsnGlnProAspValArgProGln 440
DB 1261 GCGCCCTGACCTGCAGCCCCCGAGCTGATATGTGAACCCAGCCAGATGTTGGGCCCCAG 1320
QY 441 ProProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlu 460
DB 1321 CCGCTTTCGCGCCGAGAGGGCCCTCTGCTGTGCTGCCGACCTGTGCTGTGCTGTGAA 1380
QY 461 ArgAlaLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly 480
DB 1381 AGGCGCAAGACTCTCTCCCGAGGGAAGATGGGGTCTCAAAAGACGTTTTTGGCTTTGGG 1440
QY 481 GlyAlaValGluAsnProGluTyrrLeuThrProGlnGlyGlyAlaAlaProGlnProHis 500
DB 1441 GGTGCGCTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGCTGCCCCCTCAGCCCCAC 1500
QY 501 ProProProAlaPheSerProAlaPheAspAsnLeuTyrrTrpAspGlnAspProPro 520
DB 1501 CTTCTCTCTGCTTTCAGCCCGAGCTTCGACAACTCTATTACTTGGGAGCCAGACCCACCA 1560
QY 521 GluArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrr 540
DB 1561 GAGCGGGGGCTCCACCCAGACCTTCAAGAGGACACTTACGCGAGAAACCCAGAGTAC 1620
QY 541 LeuGlyLeuAspValProAlaAlaAlaProAlaArgSerProSerProSerThrGlnPro 560
DB 1621 CTGGGTCTGGAGCTGGCGCGCGCACCCCGCTCGCCCGAGCCCGAGCACACAGCC 1680
QY 561 TrpGluHisValAsnAlaIleGlnGluAlaArgArgLeuLeuAsnLeuSerArgAspThr 580
DB 1681 TGGGAGCATGTGAATGCCATCCAGGAGCGCGGCTCTCTCTGAACTCTGAGTAGAGACT 1740
QY 581 AlaAlaGluMetAsnGluThrValGluValIleSerGluMetPheAspLeuGlnGluPro 600
DB 1741 GCTGCTGAGATGAATGAACAGTAGAAGTATCTCAGAAATGTTTGACTCCAGAGCCG 1800
QY 601 ThrCysLeuGlnThrArgLeuGluLeuTyrrLysGlnGlyLeuArgGlySerLeuThrLys 620
DB 1801 ACCTGCTACAGACCCCGCTGGAGCTGTACAAGCAGGCGCTCGCGGGCAGCCTCACCAAG 1860
QY 621 LeuLysGlyProLeuThrMetMetAlaSerHisTyrrLysGlnHisCysProProThrPro 640
DB 1861 CTCAGGCGCCCTTGGACCATGTGCGCAGCCACTACAAACAGACGACTGCGCTCCCAACCCG 1920
QY 641 GluThrSerCysAlaThrGlnIleIleThrPheGluSerPheLysGluAsnLeuLysAsp 660
DB 1921 GAAACTCTCTGTGACCCAGATTTACCTTTGAAAGTTTCAAGAGAACCTGAGAGAC 1980
QY 661 PheLeuLeuValIleProPheAspCysTrpGluProValGlnGluGlyAlaProProPro 680

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2423	64.8	4530	1	US-08-229-515A-9	Sequence 9, Appli
2	2423	64.8	4530	1	US-08-645-865-9	Sequence 9, Appli
3	2423	64.8	4530	4	US-09-167-322-4	Sequence 4, Appli
4	2423	64.8	4530	4	US-09-527-487-1	Sequence 1, Appli
5	2423	64.8	4530	4	US-09-877-177A-11	Sequence 11, Appli
6	2422	64.8	9374	4	US-09-811-115-1	Sequence 1, Appli
7	2418	64.7	4473	2	US-09-048-804-1	Sequence 1, Appli
8	2418	64.7	4473	3	US-09-056-105-26	Sequence 26, Appli
9	2418	64.7	4473	4	US-09-663-834A-3	Sequence 3, Appli
10	2410.5	64.5	3768	4	US-09-811-115-2	Sequence 2, Appli
11	2405.5	64.3	3768	2	US-08-625-101-1	Sequence 1, Appli
12	2405.5	64.3	3768	2	US-08-356-786-1	Sequence 1, Appli

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:04:16 ; Search time 679.537 Seconds
(without alignments)
4313.608 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739

Sequence: 1 MRAAPLLARASLSGLF.....EPVQEGAPPPPAAHHHHH 690

Scoring table:

	BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0			0.5
Fgapop 6.0			7.0
Delop 6.0			7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO/spool/US9821883/runat_09092004_105126_6623/app_query.fasta_1.3100
-DB=US_Geneseq_29Jan04 -QFMT=fastacp -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US9821883 @CGN_1_1646 @runat_09092004_105126_6623 -NCPU=6 -ICPU=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

No.	Score	Query	Match	Length	ID	Description
1	3739	100.0	2070	4	AAD21565	AAD21565 Human HER
2	3473.5	92.9	2091	4	AAD21567	AAD21567 Human HER
3	2974	79.5	1665	4	AAD21564	AAD21564 Human HER
4	2959.5	79.2	1692	4	AAD21566	AAD21566 Human HER
5	2423	64.8	4530	2	AAT01585	AAT01585 Her-2/neu
6	2423	64.8	4530	2	AAT71253	AAT71253 Human HER
7	2423	64.8	4530	3	AAD260815	AAD260815 Nucleotid
8	2423	64.8	4530	4	AAD19731	AAD19731 Human tyr

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	ID	Description
1	3739	100.0	2070	4	AAD21565	AAD21565 Human HER
2	3473.5	92.9	2091	4	AAD21567	AAD21567 Human HER
3	2974	79.5	1665	4	AAD21564	AAD21564 Human HER
4	2959.5	79.2	1692	4	AAD21566	AAD21566 Human HER
5	2423	64.8	4530	2	AAT01585	AAT01585 Her-2/neu
6	2423	64.8	4530	2	AAT71253	AAT71253 Human HER
7	2423	64.8	4530	3	AAD260815	AAD260815 Nucleotid
8	2423	64.8	4530	4	AAD19731	AAD19731 Human tyr

ID	Score	Query	Match	Length	ID	Description
9	2423	64.8	4530	6	ABN85585	ABN85585 Human HER
10	2423	64.8	4530	6	ABZ35012	ABZ35012 Human gen
11	2423	64.8	4530	6	ABV94128	ABV94128 Breast ca
12	2423	64.8	4530	6	ABK83918	ABK83918 Human cDN
13	2423	64.8	4530	7	ACC50139	ACC50139 Breast ca
14	2423	64.8	4530	7	ABQ83856	ABQ83856 Human Her
15	2423	64.8	4530	8	AAD58073	AAD58073 Human C-e
16	2423	64.8	4530	9	ADC09594	ADC09594 Her2/Neu
17	2422	64.8	9274	6	AAD43934	AAD43934 HER-2 tra
18	2422	64.8	9274	6	ABK14057	ABK14057 Human HER
19	2418	64.7	4472	3	AA14812	AA14812 CDNA enco
20	2418	64.7	4472	3	ABQ76220	ABQ76220 Human tum
21	2418	64.7	4473	2	AAZ31071	AAZ31071 HER-2 nuc
22	2418	64.7	4473	6	ABZ34969	ABZ34969 Human gen
23	2418	64.7	4473	6	AAZ38904	AAZ38904 Human Her
24	2418	64.7	4473	7	ACC69999	ACC69999 Human C-e
25	2418	64.7	4473	9	ADC35148	ADC35148 Human bre
26	2418	64.7	4473	9	ADD25483	ADD25483 Binding d
27	2415	64.6	9274	4	AAF24297	AAF24297 HER2 tran
28	2410.5	64.5	3765	9	ADB67620	ADB67620 Human epi
29	2410.5	64.5	3768	3	AAA09455	AAA09455 Human her
30	2410.5	64.5	3768	6	ABV78168	ABV78168 Human ERB
31	2410.5	64.5	3768	6	ABZ35744	ABZ35744 Human ERB
32	2410.5	64.5	3768	6	AAZ43935	AAZ43935 Human HER
33	2410.5	64.5	3768	6	ABX09987	ABX09987 Human ERB
34	2410.5	64.5	3768	6	AAZ43986	AAZ43986 Human Her
35	2410.5	64.5	3768	6	ABK14058	ABK14058 Human HER
36	2410.5	64.5	3768	6	ABL91709	ABL91709 Human pol
37	2410.5	64.5	3768	7	ACC57649	ACC57649 Human pro
38	2405.5	64.3	3768	2	AAT40739	AAT40739 HER-2/neu
39	2405.5	64.3	3768	2	AAZ01912	AAZ01912 Human HER
40	2405.5	64.3	3768	5	AAH23392	AAH23392 Human HER
41	2405.5	64.3	3768	6	AAZ32743	AAZ32743 Human Her
42	2405.5	64.3	3768	6	ABK10730	ABK10730 Human Her
43	2405.5	64.3	3768	6	ABA92250	ABA92250 Human Her
44	2396.5	64.1	4299	2	AAZ46083	AAZ46083 Sequence
45	2393	64.0	4606	9	ADB47370	ADB47370 Human cDN

ALIGNMENTS

RESULT 1

AAD21565
ID AAD21565 standard; DNA; 2070 BP.

XX AC AAD21565;

XX DT 28-JAN-2002 (first entry)

XX DE Human HER500-hGM-CSF fusion DNA construct.

XX KW Immunostimulatory fusion protein; IFP; antigen component; therapy;
KW immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
KW HER-2 protein; granulocyte-macrophage colony stimulating factor;
KW HER500-hGM-CSF fusion DNA; ds.

XX OS Homo sapiens.

OS Synthetic.

OS Chimeric.

XX FH Key Location/Qualifiers

XX CDS 1..2070

FT /product= "Human HER500-hGM-CSF fusion protein construct
FT comprising human PAP signal sequence, mature PAP protein,
FT an Ala Arg linker, human HER-2 signal sequence, mature
FT HER-2 membrane distal extracellular and intracellular
FT domains, an Ala Ala linker, a mature human GM-CSF
FT sequence and a C-terminal tag"
FT /note= "CDS does not include stop codon"

```

FT XX /partial
PN XX WO200174855-A2.
XX XX
PD XX 11-OCT-2001.
XX XX
PF XX 30-MAR-2001; 2001WO-US010515.
XX XX
PR XX 30-MAR-2000; 2000US-0193504P.
XX XX
XX XX (DEND-) DENDREON CORP.
XX XX
XX XX Laus R, Vidovic D, Graddis T;
XX XX
DR WPI; 2001-662965/76.
XX XX
XX XX P-PSDB; AAE13109.
XX XX

XX XX An immunostimulatory fusion protein comprising the intracellular domain
XX XX of HER-2 and an antigen elicits an immune response to the antigen and is
XX XX useful for the treatment of associated cancer associated.
XX XX
XX XX Example 3; Page 28; 59pp; English.
XX XX
XX XX The invention relates to immunostimulatory fusion proteins (IPF) and
XX XX nucleic acid molecules encoding such proteins. The IPFs comprise a
XX XX polypeptide antigen component and an immunostimulatory component derived
XX XX from the intracellular domain of HER-2 protein which is effective to
XX XX elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
XX XX immune response to the antigen. IPF or superactivated dendritic cells are
XX XX used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
XX XX associated with a particularly antigen. The present sequence is HER500
XX XX hGM-CSF fusion DNA construct which comprises DNA molecules encoding human
XX XX PAP signal sequence, mature PAP protein, an Ala Arg linker, human HER-2
XX XX signal sequence, mature HER-2 membrane distal extracellular and
XX XX intracellular domains, an Ala Ala linker, a mature human granulocyte-
XX XX macrophage colony stimulating factor (GM-CSF) sequence and a C-terminal
XX XX tag
XX XX
XX XX Sequence 2070 BP; 425 A; 707 C; 568 G; 370 T; 0 U; 0 Other;
XX XX
XX XX Alignment Scores:
XX XX Pred. No.: 1.89e-154 Length: 2070
XX XX Score: 3739.00 Matches: 690
XX XX Percent Similarity: 100.00% Conservative: 0
XX XX Best Local Similarity: 100.00% Mismatches: 0
XX XX Query Match: 100.00% Indels: 0
XX XX DB: 4 Gaps: 0
XX XX
XX XX US-09-821-883-2 (1-690) x AAD21565 (1-2070)
XX XX
QY 1 MetArgAlaAlaProLeuLeuLeuAlaAlaArgAlaAlaSerLeuSerLeuGlyPheIeuPhe 20
DB 1 ATGAGAGTGACCCCTCTCTCGGCCAGGCGACGAGCCCTTAGCCCTTGGCTTCTTGTTT 60
QY 21 LeuLeuPheThrLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAla 40
DB 61 CTGCTTTTTTCTGGCTAGACCGAAGTGATACGCCAAGAGGTTGGCGCGCGGCGCG 120
QY 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
DB 121 TCGACCCCAAGTGTGCACCGGCACACATGAGCTGGCGCTCCCTGGCCAGTCCCGAGACC 180
QY 61 HisLeuAspMetLeuArgHisLeuThrGlnGlyCysGlnValValGlnGlyAsnLeuGlu 80
DB 181 CACCTGGACATGCTCCCGCACCTCTACACAGGCGTGCAGGTGGTGCAGGGAACCTGGAA 240
QY 81 LeuThrThrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGln 100
DB 241 CTACCTACCTGCGCCCAACCAATGCCAGCTGCTCTCTCTGCAAGATATCCAGAGAGGTGCAG 300
QY 101 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
DB 301 GGCTACGTGCTCATCGCTCAACCAAGTGAGGCGAGTCCCACTGCAGAGGCTCGGATT 360

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:41:11 ; Search time 7392.31 Seconds
(without alignments)
4045.651 Million cell updates/sec

Title: US-09-821-883-2
Perfect score: 3739
Sequence: 1 MRAAPLLARAASLSGLF.....EPVQEGAPPPAAHHHHH 690

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09821883/runat.09092004.105127.6631/app.query.fasta.1.3100
-DB=genEmbl -QMT=frascap -SUFFIX=rge -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09821883 @CGN 1.1 16795 @runat.09092004.105127.6631 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_ats:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_nu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_ats:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3739	100.0	2070	6	AX268286 Sequence
2	3473.5	92.9	2091	6	AX268288 Sequence
3	2974	79.5	1665	6	AX268285 Sequence
4	2959.5	79.2	1692	6	AX268287 Sequence
5	2423	64.8	4530	6	I21124 Sequence 9
6	2423	64.8	4530	6	I59745 Sequence 9
7	2423	64.8	4530	6	AR202597 Sequence
8	2423	64.8	4530	6	AR283481 Sequence
9	2423	64.8	4530	6	AR344811 Sequence
10	2423	64.8	4530	6	AX282577 Sequence
11	2423	64.8	4530	6	AX587649 Sequence
12	2423	64.8	4530	6	AX644071 Sequence
13	2423	64.8	4530	6	AX771418 Sequence
14	2423	64.8	4530	6	BD005474 Cellular
15	2423	64.8	4530	9	HUMHER2A
16	2422	64.8	9274	6	AR409602 Sequence
17	2422	64.8	9274	6	AX060703 Sequence
18	2418	64.7	4473	6	AR080259 Sequence
19	2418	64.7	4473	6	AR167390 Sequence
20	2418	64.7	4473	6	AR392088 Sequence
21	2418	64.7	4473	9	HSERB2R
22	2410.5	64.5	3768	6	AR409603 Sequence
23	2410.5	64.5	3768	6	AX060704 Sequence
24	2410.5	64.5	3768	6	AX467229 Sequence
25	2410.5	64.5	3768	6	AX481438 Sequence
26	2410.5	64.5	3768	6	BD224136 Novel met
27	2405.5	64.3	3768	6	AR034479 Sequence
28	2405.5	64.3	3768	6	BD267514 HER-2/neu
29	2405.5	64.3	3768	6	AX201817 Sequence
30	2405.5	64.3	3768	6	AX380923 Sequence
31	2405.5	64.3	3768	6	AX384604 Sequence
32	2405.5	64.3	3768	6	AX465456 Sequence
33	2334.5	62.4	3678	6	AX505114 Sequence
34	2258.5	60.4	2385	6	AR082744 Sequence
35	2258.5	60.4	2385	6	AR099963 Sequence
36	2258.5	60.4	2385	6	AR143949 Sequence
37	2188.5	58.5	1437	6	AX268289 Sequence
38	2130	57.0	2781	6	AX380944 Sequence
39	2129	56.9	2763	6	AX380942 Sequence
40	2124.5	56.8	3780	4	AB008451 Canis fam
41	1974	52.8	4506	10	AK129487 Mus muscu
42	1972	52.7	4727	10	AY116182 Rattus no
43	1968	52.6	3955	6	BD267515 HER-2/neu
44	1968	52.6	3955	6	I21129 Sequence 14
45	1968	52.6	3955	6	I59750 Sequence 14

ALIGNMENTS

AX268286
LOCUS AX268286 2070 bp DNA linear PAT 29-OCT-2001
DEFINITION Sequence 7 from Patent WO0174855.
ACCESSION AX268286
VERSION AX268286.1 GI:16541538
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE
1
AUTHORS Laus, R., Vidovic, D. and Graddis, T.
TITLE Compositions and methods for dendritic cell-based immunotherapy
JOURNAL Patent: WO 0174855-A 7 11-OCT-2001;
DENDREON CORPORATION (US)
FEATURES
source Location/Qualifiers
1..2070
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="HER500-hgm-CSF construct"

ORIGIN
Alignment Scores:
Pred. No.: 7,45e-129 Length: 2070
Score: 3739,00 Matches: 690
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-821-883-2 (1-690) x AX268286 (1-2070)

Qy 1 MetArgAlaProLeuLeuLeuAlaArgAlaSerLeuSerLeuGlyPheLeuPhe 20
Db 1 ATGAGAGCTGCACCCCTCCTCTGGCCAGGACAGCAAGCCCTTAGCCCTTGGCTTCTGTTT 60
Qy 21 LeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGlyLeuLeuAlaArgGlyAlaAla 40
Db 61 CTGCTTTTCTGGCTAGACCCGAAAGTGTTACTACCAAGAGAGTTGGCGCGGCGCGCG 120
Qy 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
Db 121 TCGACCCCAAGTGTGCACCGGCACAGACATGAAGCTGGCGCTCCCTGCGCAGTCCCGAGACC 180
Qy 61 HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeuGlu 80
Db 181 CACCTGGACATGTCTCCGCCACCTCTACACGGGCTGCCAGGTGGTGCAGGGGAAACCTGGAA 240
Qy 81 LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGln 100
Db 241 CTACCTACCTGCACCAACATGCCAGCTGCTCTCTCTGCGGATATCCAGGAGTGCAG 300
Qy 101 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
Db 301 GGCTACGTGCTCATCGCTCAACCAAGTAGGAGGAGTCCCACTGCAGAGGCTCGCGATT 360
Qy 121 ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 140
Db 361 GTGCGAGGCCACCCAGCTCTTTGAGGACAACTATATCCCTGGCCGTGTAGACAATGGAGAC 420
Qy 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln 160
Db 421 CCGCTGAACATACACCCCTGTCTACAGGGGCTCCCGAGGAGGCTTCGGGGAGCTGCAG 480
Qy 161 LeuArgSerLeuThrGluLeuLeuLysGlyGlyValLeuIleGlnArgAsnProGlnLeu 180
Db 481 CTTGGAAGCTTCACAGAGATCTTGAAGAGGGGCTTGTATCCAGCGCAACCCCGAGCTC 540
Qy 181 CysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeuLeuLeu 200
Db 541 TGCTACGAGGACAGATTTGTGGAGAGACATCTCCACAGAAACAACAGCTGGCTCTC 600
Qy 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220

Db 601 ACACGTAGACACCAACCGCTCTCGGGCTCGCCACCCCTGTTCTCCGATGTGTAAGGCG 660
Qy 221 SerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAla 240
Db 661 TCCCGCTGCTGGGAGAGAGTTCTGAGGATTTGTGAGAGCCCTGACGCCACTGTCTGTGCC 720
Qy 241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCysAla 260
Db 721 GGTGGCTGTGCCCTGCAAGGGGCCACTGCGCCACTGCTGCTGCTCATGAGCATGTGCT 780
Qy 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
Db 781 GCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGCTGCCCTCCACTTCAACACAGT 840
Qy 281 GlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSer 300
Db 841 GGCATCTGTGAGTGTGACTGCCAGCCCTGGTACCTTACCAACACAGACACGCTTTGAGTCC 900
Qy 301 MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr 320
Db 901 ATGCCCAATCCGAGGCGCGGTATACATTCGGGCCAGCTGTGTGACTGCTGCTGCCCTAC 960
Qy 321 AsnTyrLeuSerThrAspValGlySerGlyValaGlyGlyMetValHisHisArgHisArg 340
Db 961 AACTACCTTTCTACGGAGCTGGGATCGGCGCTGGGGGCATGGTCCACACAGCACCGC 1020
Qy 341 SerSerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGlu 360
Db 1021 AGCTCATCTACAGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAG 1080
Qy 361 GluAlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGly 380
Db 1081 GAGGCCCCAGAGTCTCCACTGGACCCCTCCGAGGGGCTGGCTCCGATGATTTGATGGT 1140
Qy 381 AspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerPro 400
Db 1141 GACCTGGGAATGGGGCAGCAAGGGCTGCAAAAGCTCCCAACACACATGATGATGCTGTT 1200
Qy 401 LeuGlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrVal 420
Db 1201 CTACAGCGTACAGTACAGGACCCCAACATGAGTACCTGCTGCTGCTGCTGCTGCTGCT 1260
Qy 421 AlaProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGln 440
Db 1261 GCCCCCTGACTGACGCCCCAGCTGATATGTGAACCAACAGCAGATGTTGCGCCCGCAG 1320
Qy 441 ProProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlu 460
Db 1321 CCCCCTTGGCCCGAGAGGGCCCTCTGCTGCTGCCGACCTGCTGCTGCTGCTGCTGCTGCT 1380
Qy 461 ArgAlaLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly 480
Db 1381 AGGGCCAAAGACTCTCTCCCGAGGGAAGATGGGGTCTGCAAGAGCTTTTGGCTTTGGG 1440
Qy 481 GlyAlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHis 500
Db 1441 GGTGCGCTGGAGAACCCCGAGTACTTGACACCCCGAGGAGAGCTGCCCCCTCAGCCCCAC 1500
Qy 501 ProProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProPro 520
Db 1501 CCTCCTCTGCTTTCAGCCCGCCCTTCGACAACTCTATTACTGGGACAGGACCCACCA 1560
Qy 521 GluArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyr 540
Db 1561 GAGCGGGGGCTCCACCGACACTTCAAGGGGACACTACGGCAGAGAACCCAGAGTAC 1620
Qy 541 LeuGlyLeuAspValProAlaAlaProAlaArgSerProSerProSerThrGlnPro 560
Db 1621 CTGGGTCTGACGTGCGCGCGCGCACCCCGCGCTCGCCCGCCCGCCAGCACACAGCCC 1680
Qy 561 TrpGluHisValAsnAlaIleGlnGluAlaArgArgLeuLeuAsnLeuSerArgAspThr 580

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:28:07 ; Search time 68.5595 Seconds
(without alignments)
3175.456 Million cell updates/sec

Title: US-09-821-883-2
Perfect score: 3739
Sequence: 1 MRAAPLLAARASLSGLF.....EPVOGAPPPPAHHHHH 690

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2124.5	56.8	1259	6	O18735 canis famil
2	1951	52.2	1259	11	Q8K3F9
3	1609.5	43.0	419	4	Q8UKY9
4	1376.5	36.8	711	11	Q80V89
5	1005.5	26.9	881	11	Q8C0E7
6	969	25.9	367	11	Q8R2X1
7	906	24.2	165	4	Q14256
8	900.5	24.1	412	4	Q8WVU0
9	861.5	23.0	431	13	Q7SYU9
10	754.5	20.2	527	13	Q90836
11	723	19.3	149	6	Q8BG66
12	717.5	19.2	643	11	Q9ERV6
13	717.5	19.2	655	11	Q9WVF5
14	717.5	19.2	1210	11	Q9EP98
15	714.5	19.1	1209	11	Q9QX70
16	708.5	18.9	1209	6	Q8MIL8

17	704.5	18.8	478	11	Q9ESE0
18	703	18.8	1191	13	Q7SZF7
19	697	18.6	331	4	Q9BUD7
20	664.5	17.8	1165	13	Q9YH40
21	657.5	17.6	1305	13	Q8AW81
22	654	17.5	144	6	Q9GL44
23	641.5	17.2	599	13	Q9PSH2
24	634	17.0	144	6	Q86SY5
25	608	16.3	1328	13	P79754
26	556	14.9	1377	5	Q8MLW0
27	545	14.6	144	6	Q9MYK4
28	543	14.5	1433	5	Q9BIH9
29	537	14.4	1322	5	Q86NZ2
30	527	14.1	146	6	Q8WNL7
31	515	13.8	152	6	Q9SLI0
32	473	12.7	141	6	Q7YRF7
33	427	11.4	141	11	Q8VH40
34	407	10.9	1137	13	Q9WGF6
35	398	10.6	150	6	Q9BG64
36	395	10.6	151	6	Q9BG65
37	382.5	10.2	145	11	Q8CFB5
38	377	10.1	1368	5	Q23821
39	362	9.7	138	11	Q9UJ91
40	341	9.1	365	5	Q26589
41	341	9.1	1717	5	Q26586
42	331	8.9	334	5	Q26567
43	331	8.9	342	5	Q26568
44	323.5	8.7	1193	5	Q9Y1X8
45	305	8.2	1564	5	Q86MD7

ALIGNMENTS

RESULT 1

O18735		PRELIMINARY;	PRT; 1259 AA.
ID	O18735		
AC	O18735;		
DT	01-JAN-1998 (TREMBLrel. 05, Created)		
DT	01-JAN-1998 (TREMBLrel. 05, Last sequence update)		
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)		
DE	Erbb-2		
OS	Canis familiaris (Dog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.		
OX	NCBI_taxid=9615;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Yokota H.;		
RT	"CDNA cloning of erbb-2 from canine mammary gland.";		
RL	Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AB008451; BAA23127.1; -		
DR	HSSP; P11362; 1FGK.		
DR	GO; GO:0016020; C:membrane; IEA.		
DR	GO; GO:0005524; F:ATP binding; IEA.		
DR	GO; GO:0005509; F:calcium ion binding; IEA.		
DR	GO; GO:0005006; F:epidermal growth factor activity; IEA.		
DR	GO; GO:0016740; F:transferase activity; IEA.		
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.		
DR	GO; GO:0007189; P:transmembrane receptor protein tyrosine kin. . .; IEA.		
DR	InterPro; IPR002048; EF-hand.		
DR	InterPro; IPR006211; Furin-like.		
DR	InterPro; IPR000494; EGFR_L_domain.		
DR	InterPro; IPR006212; Furin repeat.		
DR	InterPro; IPR009030; Grow_fac_recep.		
DR	InterPro; IPR000719; Prot_kinase.		
DR	InterPro; IPR001245; Tyr_kinase.		
DR	InterPro; IPR008266; Tyr_kinase_AS.		
DR	InterPro; IPR004019; YLP motif.		
DR	Pfam; PF00757; Furin-like; 1.		
DR	Pfam; PF00069; kinase; 1.		
DR	Pfam; PF01030; Recep_L_domain; 2.		
DR	Pfam; PF02757; YLP; 2.		

Q9ese0	rattus norv
Q7szf7	brachydanio
Q9bud7	homo sapien
Q9yh40	xiphophorus
Q8aw81	brachydanio
Q9gl44	macaca mula
Q9psb2	gallus gall
Q86sy5	papio anbi
P79754	fugu rubrip
Q8mlw0	drosophila
Q9myk4	ovis aries
Q9bih9	anopheles g
Q86nz2	drosophila
Q8wnl7	equus cabal
Q9sl10	equus cabal
Q7yrf7	felis silve
Q8vh40	sigmodon hi
Q9wgf6	gallus gall
Q9bg64	oryctolagus
Q9bg65	oryctolagus
Q8cfb5	meriones un
Q23821	caenorhabdi
Q9uj91	marmota mon
Q26589	schistosoma
Q26586	schistosoma
Q26567	schistosoma
Q26568	schistosoma
Q9y1x8	ephydatia f
Q86md7	echinococcu

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:20:54 ; Search time 12.6656 Seconds
(without alignments)
2836.691 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739

Sequence: 1 MRAAPLLARAASLSGLF.....EPVQEGAPPPPAHHHHH 690

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2405.5	64.3	1255	1 ERB2_HUMAN	P04626 homo sapien
2	1947.5	52.1	1257	1 ERB2_RAT	P06494 rattus norv
3	1935	51.8	1254	1 ERB2_MESAU	Q60553 mesocricetu
4	755.5	20.2	703	1 EGFR_CHICK	P13387 gallus gall
5	721	19.3	1210	1 EGFR_HUMAN	P00533 homo sapien
6	721	19.3	1308	1 ERB4_HUMAN	Q15303 homo sapien
7	717.5	19.2	1210	1 EGFR_MOUSE	Q01279 mus musculu
8	697	19.2	1308	1 ERB4_RAT	Q62956 rattus norv
9	697	18.6	1342	1 ERB3_HUMAN	P21860 homo sapien
10	674	18.0	144	1 CSF2_HUMAN	P04141 homo sapien
11	668	17.9	1339	1 ERB3_RAT	Q62799 rattus norv
12	655.5	17.5	1167	1 XMRK_XIPMA	P13388 xiphophorus
13	549	14.7	144	1 CSF2_SHEEP	P28773 ovis aries
14	536	14.3	1426	1 EGFR_DROME	P04412 drosophila
15	529	14.1	144	1 CSF2_CEREL	P51748 cervus elap
16	488	13.1	144	1 CSF2_PIG	Q29118 sus scrofa
17	481.5	12.9	143	1 CSF2_BOVIN	P11052 bos taurus
18	476.5	12.7	144	1 CSF2_CANFA	P48749 canis famil
19	452	12.1	144	1 CSF2_FELCA	Q62757 felis silve
20	441	11.8	127	1 CSF2_RAT	P48750 rattus norv
21	435.5	11.6	140	1 CSF2_CAVPO	Q60481 cavia porce
22	404.5	10.8	1367	1 LTJ3_CAEEL	P24348 caenorhabdi
23	367	9.8	141	1 CSF2_MOUSE	P01587 mus musculu
24	317.5	8.5	1363	1 ILPR_BRALA	Q02466 brachiosco
25	293	7.8	145	1 ERB2_MOUSE	P70424 mus musculu
26	273.5	7.3	1477	1 HTK7_HYDAT	Q25197 hydra atten
27	270.5	7.2	2146	1 INSR_DROME	P09208 drosophila
28	256.5	6.9	1382	1 INSR_HUMAN	P06213 homo sapien
29	252.5	6.8	1372	1 INSR_MOUSE	P15208 mus musculu
30	252	6.7	1607	1 MIPR_LYMST	Q25410 lymanaea sca
31	251.5	6.7	1300	1 IRR_MOUSE	Q9wt14 mus musculu
32	249	6.7	1383	1 INSR_RAT	P15127 rattus norv
33	248.5	6.6	1297	1 IRR_HUMAN	P14616 homo sapien

34	246.5	6.6	1300	1 IRR_CAVPO	P14617 cavia porce
35	246	6.6	1367	1 IGR_HUMAN	P08069 homo sapien
36	235.5	6.3	1370	1 IGR_RAT	P24062 rattus norv
37	233.5	6.2	1373	1 IGR_MOUSE	Q60751 mus musculu
38	228.5	6.1	581	1 IRR_RAT	Q64716 rattus norv
39	207	5.5	1390	1 INSR_AEDAE	Q93105 aedes aegypt
40	178	4.8	2715	1 MLL4_HUMAN	Q9um06 homo sapien
41	168	4.5	386	1 PPAP_HUMAN	P15309 homo sapien
42	166	4.4	707	1 SPFQ_HUMAN	P23246 homo sapien
43	166	4.4	5262	1 MLL2_HUMAN	O14686 homo sapien
44	162	4.3	1321	1 IRS2_MOUSE	P81122 mus musculu
45	160	4.3	830	1 SREC_HUMAN	Q14162 homo sapien

ALIGNMENTS

RESULT 1
ERB2_HUMAN
ID ERB2_HUMAN STANDARD; PRT; 1255 AA.
AC P04626;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell surface receptor HER2) (MLN 19).
DE ERBB2 OR HER2 OR NGL OR NEU.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86119663; PubMed=3003577;
RA Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N., Saito T., Toyoshima K.;
RT "Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth factor receptor.";
RL Nature 319:230-234(1986).
[2]
RP SEQUENCE FROM N.A., AND VARIANT ALA-1170.
RX MEDLINE=86070181; PubMed=2999974;
RA Coussens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A., McGrath J., Seeburg P.H., Libermann T.A., Schlessinger J., Francke U., Levinson A., Ullrich A.;
RT "Tyrosine kinase receptor with extensive homology to EGF receptor shares chromosomal location with neu oncogene.";
RL Science 230:1132-1139(1985).
[3]
RP SEQUENCE FROM N.A., AND VARIANTS CYS-452; VAL-655 AND ALA-1170.
RX Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M.-W., Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D., Schackwitz W.S., Sherwood J.K., Witlak L.A., Nickerson D.A.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE OF 737-1031 FROM N.A.
RX MEDLINE=8617629; PubMed=295967;
RA Semba K., Kamata N., Toyoshima K., Yamamoto T.;
RT "A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epidermal growth factor-receptor gene and is amplified in a human salivary gland adenocarcinoma.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).
[5]
RP VARIANTS VAL-654 AND VAL-655.
RX MEDLINE=93194196; PubMed=8095488;
RA Eneani A., Low J., Wallace R.B., Wu A.M.;
RT "Characterization of a new allele of the human ERBB2 gene by allele-specific competition hybridization.";
RL Genomics 15:426-429(1993).
CC -!- FUNCTION: Essential component of a neueregulin-receptor complex, although neuregulins do not interact with it alone. GP30 is a potential ligand for this receptor. Not activated by EGF, TGF-

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:29:10 ; Search time 22.0271 Seconds
(without alignments)
3013.200 Million cell updates/sec

Title: US-09-821-883-2
Perfect score: 3739
Sequence: 1 MRAAPLLARAASLSGLF.....BFVOEGAPPPAAHHHHH 690

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2405.5	64.3	1255	A24571	protein-tyrosine k
2	1950.5	52.2	1260	TVRTNU	protein-tyrosine k
3	1935	51.8	1254	I48161	p-185 precursor -
4	755.5	20.2	1223	TVCHLV	epidermal growth f
5	754.5	20.2	527	A42032	epidermal growth f
6	721	19.3	1210	GQHUE	epidermal growth f
7	721	19.3	1308	A47253	epidermal growth f
8	717.5	19.2	1210	A53183	epidermal growth f
9	714.5	19.1	544	A36325	epidermal growth f
10	697	18.6	1342	A36223	kinase-related tra
11	674	18.0	144	FQHGM	granulocyte-macrop
12	662	17.7	1339	JC4387	epidermal growth f
13	655.5	17.5	1166	S06142	protein-tyrosine k
14	549	14.7	144	JH0469	granulocyte-macrop
15	545	14.6	144	A61632	granulocyte-macrop
16	536	14.3	843	A27131	epidermal growth f
17	481.5	12.9	143	FQBOGM	granulocyte-macrop
18	476.5	12.7	144	A44936	granulocyte-macrop
19	441	11.8	127	I46269	granulocyte-macrop
20	404.5	10.8	1323	E88257	protein let-23 (im
21	404.5	10.8	1374	S70712	protein-tyrosine k
22	390.5	10.4	1369	S70713	protein-tyrosine k
23	375.5	10.0	1330	GQFPE	epidermal growth f
24	374	10.0	153	FQMSGM	granulocyte-macrop
25	341	9.1	366	D45558	epidermal growth f
26	341	9.1	1717	A45558	epidermal growth f
27	331	8.9	333	B45558	epidermal growth f
28	331	8.9	342	C45558	epidermal growth f
29	317.5	8.5	1363	T43220	insulin-like growth

30	273.5	7.3	1477	2	T18534	protein-tyrosine k
31	270.5	7.2	2101	2	S57245	insulin receptor (
32	270.5	7.2	2148	1	A56081	insulin receptor -
33	256.5	6.9	1382	1	INHUR	insulin receptor p
34	252.5	6.8	1372	2	A34157	insulin-like growth
35	252	6.7	1607	2	T43212	insulin-like growth
36	249	6.7	1383	2	A36080	insulin receptor p
37	246.5	6.6	1300	2	A36502	insulin receptor-r
38	246	6.6	1367	1	IGHUR1	insulin-like growth
39	238	6.4	540	2	B47417	insulin receptor-r
40	235.5	6.3	183	2	JH0803	tyrosine kinase re
41	235.5	6.3	1371	2	A33837	insulin-like growth
42	231.5	6.2	1268	2	B36502	insulin receptor-r
43	223.5	6.0	329	2	A48805	insulin-like growth
44	207	5.5	1390	2	T30346	insulin receptor -
45	188	5.0	1846	2	T42047	insulin receptor h

ALIGNMENTS

RESULT 1

A24571

protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human

N:Alternate names: c-erb-B-2 protein precursor; kinase-related transforming protein erbB

C:Species: Homo sapiens (man)

C>Date: 25-Oct-1987 #sequence_revision 06-Dec-1996 #text_change 11-Jun-1999

C:Accession: A24571; A25491; A44188; B44188; I59509; I57622

R:Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T.; Nature 319, 230-234, 1986

A:Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth f

A:Reference number: A24571; MUID:86118663; PMID:3003577

A:Accession: A24571

A:Molecule type: mRNA

A:Residues: 1-1255 <YAM>

A:CROSS-references: GB:X03363; NID:G31197; PIDN:CAA27060.1; PID:G31198

R:Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T. Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985

A:Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epiderm

A:Reference number: A25491; MUID:86016729; PMID:2995967

A:Accession: A25491

A:Molecule type: DNA

A:Residues: 737-1031 <SEM>

A:CROSS-references: GB:M11767; NID:G182163; PIDN:AAA35808.1; PID:G553282

R:Coussens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg, F Science 230, 1132-1139, 1985

A:Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chromo

A:Reference number: A44188; MUID:86070181; PMID:2999974

A:Accession: A44188

A:Molecule type: DNA

A:Residues: 740-910 <COU1>

A:CROSS-references: GB:M12036; NID:G183988; PIDN:AAA35978.1; PID:G183989

A:Accession: B44188

A:Molecule type: mRNA

A:Residues: 1-517; RALL, 522, 'S', 524-554, 'V', 656-1169, 'A', 1171-1255 <COU2>

A:CROSS-references: GB:M11730; NID:G183986

R:King, C.R.; Kraus, M.H.; Aaronson, S.A. Science 229, 974-976, 1985

A:Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.

A:Reference number: I59509; MUID:85272597; PMID:2992089

A:Accession: I59509

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 832-909 <REX>

A:CROSS-references: GB:L29395; NID:G459807; PIDN:AAA35809.1; PID:G459808

R:Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D. Mol. Cell. Biol. 7, 2597-2601, 1987

A:Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptional

A:Reference number: I57622; MUID:87286898; PMID:3039351

A:Accession: I57622

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-191 <TAL>


```
Db 121 VRGTLFEDNYALAVLNDGDPPLNNTTPTVGTASPGGLRELQLRSLEILKGGVLIQRNPQL 180
Qy 181 CYQDTILWKDIFHKNNQALALIDTNRSRACHPCSPMCKGRCWGSESDCQSLTRTVCA 240
Db 181 CYQDTILWKDIFHKNNQALALIDTNRSRACHPCSPMCKGRCWGSESDCQSLTRTVCA 240
Qy 241 GGCARCKGPLPTDCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPES 300
Db 241 GGCARCKGPLPTDCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPES 300
Qy 301 MPNPEGRTYTFGASCVTACPNYLSLTDVSGAGGVMVHRRHRSSTRSRGGDLTLGLPSEE 360
Db 301 MPNPEGRTYTFGASCVTACPNYLSLTDVSGAGGVMVHRRHRSSTRSRGGDLTLGLPSEE 360
Qy 361 EAPRSPLAPSGAGSDVDFDGLGMAAGLQSLPHTDPSPLQRYSEDPTVPLPSETDGYV 420
Db 361 EAPRSPLAPSGAGSDVDFDGLGMAAGLQSLPHTDPSPLQRYSEDPTVPLPSETDGYV 420
Qy 421 APLTCSPOPEYVNPQDVRPQPPSPREGPLPAARPAGATLERAKTILSPGKNGVVKDVFAG 480
Db 421 APLTCSPOPEYVNPQDVRPQPPSPREGPLPAARPAGATLERAKTILSPGKNGVVKDVFAG 480
Qy 481 GAVENPEYLTPOGGAAPOPHPPAFSPAFDNLTYWDQDPPPERGAPPSTFKGTPTAENPEY 540
Db 481 GAVENPEYLTPOGGAAPOPHPPAFSPAFDNLTYWDQDPPPERGAPPSTFKGTPTAENPEY 540
Qy 541 LGLDVPAAPARSPSPSTQPMHEVNAIOEARLLNLSRDTAAEMNETVEVISEMFDLQEP 600
Db 541 LGLDVPAAPARSPSPSTQPMHEVNAIOEARLLNLSRDTAAEMNETVEVISEMFDLQEP 600
Qy 601 TCLQTRLELYKQGLRGSITKLKGPLTMASHYKQHCPTPTSCATQIITPESFKENLKD 660
Db 601 TCLQTRLELYKQGLRGSITKLKGPLTMASHYKQHCPTPTSCATQIITPESFKENLKD 660
Qy 661 FLVLPDCEPVEQEGAPPPPPAAAHHHHHH 690
Db 661 FLVLPDCEPVEQEGAPPPPPAAAHHHHHH 690

RESULT 2
US-09-821-883-4
; Sequence 4, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; PRIOR FILING DATE: 2001-03-30
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500*-rGM-CSF construct
US-09-821-883-4

Query Match 92.9%; Score 3473.5; DB 9; Length 697;
Best Local Similarity 91.7%; Pred. No. 1.7e-226;
Matches 641; Conservative 19; Mismatches 28; Indels 11; Gaps 2;
```

```
Qy 1 MRAAPLLIARAASLSLGLFLFLFWLDRSVLAKELARGAASQTGCTGDMKRLPASPSET 60
Db 1 MRAAPLLIARAASLSLGLFLFLFWLDRSVLAKELARGAASQTGCTGDMKRLPASPSET 60
```

```
Qy 61 HLDMLRHLQCVQVQGNLELTYLPTNASLSFLQDIOEVQGYVLIAHNQVQVPLQRLRI 120
Db 61 HLDMLRHLQCVQVQGNLELTYLPTNASLSFLQDIOEVQGYVLIAHNQVQVPLQRLRI 120
Qy 121 VRGTLFEDNYALAVLNDGDPPLNNTTPTVGTASPGGLRELQLRSLEILKGGVLIQRNPQL 180
Db 121 VRGTLFEDNYALAVLNDGDPPLNNTTPTVGTASPGGLRELQLRSLEILKGGVLIQRNPQL 180
Qy 181 CYQDTILWKDIFHKNNQALALIDTNRSRACHPCSPMCKGRCWGSESDCQSLTRTVCA 240
Db 181 CYQDTILWKDIFHKNNQALALIDTNRSRACHPCSPMCKGRCWGSESDCQSLTRTVCA 240
Qy 241 GGCARCKGPLPTDCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPES 300
Db 241 GGCARCKGPLPTDCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPES 300
Qy 301 MPNPEGRTYTFGASCVTACPNYLSLTDVSGS-----GAGGMVHRRHRSSTRSRGGDL 351
Db 301 MPNPEGRTYTFGASCVTACPNYLSLTDVSGSIIINFEKLGAGGMVHRRHRSSTRSRGGDL 360
Qy 352 TLGLEPSEERAPSPPLAPSGAGSDVDFDGLGMAAGLQSLPHTDPSPLQRYSEDPTVP 411
Db 361 TLGLEPSEERAPSPPLAPSGAGSDVDFDGLGMAAGLQSLPHTDPSPLQRYSEDPTVP 420
Qy 412 LPSETDGYVAPLTCSPOPEYVNPQDVRPQPPSPREGPLPAARPAGATLERAKTILSPGKNG 471
Db 421 LPSETDGYVAPLTCSPOPEYVNPQDVRPQPPSPREGPLPAARPAGATLERAKTILSPGKNG 480
Qy 472 VVKDVFAGGAVENPEYLTPOGGAAPOPHPPAFSPAFDNLTYWDQDPPPERGAPPSTFKG 531
Db 481 VVKDVFAGGAVENPEYLTPOGGAAPOPHPPAFSPAFDNLTYWDQDPPPERGAPPSTFKG 540
Qy 532 TPTAENPEYLGIDVPAAPARSPSPSTQPMHEVNAIOEARLLNLSRDTAAEMNETVEVI 591
Db 541 TPTAENPEYLGIDVPAAPARSPSPSTQPMHEVNAIOEARLLNLSRDTAAEMNETVEVI 600
Qy 592 SEMFDLQEPDCLQTRLELYKQGLRGSITKLKGPLTMASHYKQHCPTPTSCATQIITF 651
Db 601 SNEFSIORPTCVQTRLELYKQGLRGSITKLKGPLTMASHYKQHCPTPTSCATQIITF 660
Qy 652 ESPKENLKDPLVLPDCEPVEQEGAPPPPPAAAHHHHHH 690
Db 661 EDFIKNLKGLFDIPDFDCEPVEQEGAPPPPP--AHHHHHH 697

RESULT 3
US-09-821-883-1
; Sequence 1, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; PRIOR FILING DATE: 2001-03-30
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500 construct
US-09-821-883-1

Query Match 79.5%; Score 2972; DB 9; Length 555;
Best Local Similarity 100.0%; Pred. No. 1.1e-192;
Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:31:25 ; Search time 20.0998 Seconds
(without alignments)
1772.256 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739

Sequence: 1 MRAAPLLARAASLSGLFLF.....EPVQEGAPPPAAAAHHHHH 690

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/ECTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2410.5	64.5	1255	4	US-09-527-487-2
2	2410.5	64.5	1255	4	US-09-811-115-3
3	2405.5	64.3	1255	1	US-08-467-083-68
4	2405.5	64.3	1255	1	US-08-414-417B-68
5	2405.5	64.3	1255	2	US-08-486-348A-68
6	2405.5	64.3	1255	2	US-08-625-101-2
7	2405.5	64.3	1255	2	US-08-468-545B-68
8	2405.5	64.3	1255	2	US-08-356-786-2
9	2405.5	64.3	1255	3	US-08-466-680B-68
10	2405.5	64.3	1255	4	US-09-354-533-68
11	2385.5	63.8	1255	2	US-08-484-438-8
12	2224	59.5	782	2	US-09-146-283-4
13	2224	59.5	782	3	US-08-579-823A-4
14	2224	59.5	782	3	US-09-344-195-4
15	1607	43.0	419	4	US-09-630-155-2
16	1597	42.7	624	3	US-08-422-108-1
17	1597	42.7	624	4	US-08-422-734-1
18	1183	31.6	580	1	US-08-414-417B-69
19	1183	31.6	580	2	US-08-486-348A-69
20	1183	31.6	580	2	US-08-468-545B-69
21	1183	31.6	580	3	US-08-466-680B-69
22	1183	31.6	580	4	US-09-354-533-69
23	860	23.0	166	4	US-09-648-067A-1
24	761	20.4	515	2	US-09-146-283-2
25	761	20.4	515	3	US-08-579-823A-2
26	761	20.4	515	3	US-09-344-195-2
27	721	19.3	644	1	US-08-336-708A-9

28	721	19.3	911	2	US-08-484-438-10	Sequence 10, Appl
29	721	19.3	1058	2	US-08-484-438-4	Sequence 4, Appl1
30	721	19.3	1210	2	US-08-484-438-7	Sequence 7, Appl1
31	721	19.3	1210	2	US-08-475-035-4	Sequence 4, Appl1
32	721	19.3	1308	2	US-08-484-438-2	Sequence 2, Appl1
33	704.5	18.8	478	4	US-09-570-454-2	Sequence 2, Appl1
34	704.5	18.7	478	4	US-09-867-521-2	Sequence 2, Appl1
35	698.5	18.7	1343	6	5183884-4	Patent No. 5183884
36	697	18.6	1342	1	US-07-978-895-4	Sequence 4, Appl1
37	697	18.6	1342	2	US-08-484-438-9	Sequence 9, Appl1
38	697	18.6	1342	2	US-08-473-119-4	Sequence 4, Appl1
39	697	18.6	1342	2	US-08-475-352-4	Sequence 4, Appl1
40	697	18.6	1342	4	US-09-170-699-4	Sequence 4, Appl1
41	693	18.5	274	3	US-08-469-318-144	Sequence 144, App
42	693	18.5	274	3	US-08-468-609A-144	Sequence 144, App
43	693	18.5	274	4	US-08-446-872A-144	Sequence 144, App
44	693	18.5	274	4	US-08-762-227A-144	Sequence 144, App
45	693	18.5	274	5	PCT-US95-01185-144	Sequence 144, App

ALIGNMENTS

RESULT 1

US-09-527-487-2
; Sequence 2, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126881309200
; CURRENT APPLICATION NUMBER: US/09/527,487
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-527-487-2

Query Match 64.5%; Score 2410.5; DB 4; Length 1255;
Best Local Similarity 41.2%; Pred. No. 3.7e-179;
Matches 510; Conservative 0; Mismatches 2; Indels 727; Gaps 1;

QY	35	LARGASTQVCTGDMKLRLPASPTHLDMLRHLHYGGVQGVQGNLELYLPTNALSFLQ	94
DB	16	LPPGAASQVCTGDMKLRLPASPTHLDMLRHLHYGGVQGVQGNLELYLPTNALSFLQ	75
QY	95	DIQEVQGYVLIHNNQVRQVPLQRLRIVRGTLFEDNYALAVLDNGDPLNNTTPTVTGASPG	154
DB	76	DIQEVQGYVLIHNNQVRQVPLQRLRIVRGTLFEDNYALAVLDNGDPLNNTTPTVTGASPG	135
QY	155	GLRELQLASLTELKGGVLIQGNPOLCQVDITLWKDIFHKNQLALTITIDNRSRACHPC	214
DB	136	GLRELQLASLTELKGGVLIQGNPOLCQVDITLWKDIFHKNQLALTITIDNRSRACHPC	195
QY	215	SPMKGRSCWGESSEDCQSLRTVCAGGCACRCKGLPTDCCHCCCAACTGPKHSDCLAC	274
DB	196	SPMKGRSCWGESSEDCQSLRTVCAGGCACRCKGLPTDCCHCCCAACTGPKHSDCLAC	255
QY	275	LHFHNSGICELHCPALVTYNTDTFESMPNPEGRVTFGASCVTACPYNYLSLTDVGS----	329
DB	256	LHFHNSGICELHCPALVTYNTDTFESMPNPEGRVTFGASCVTACPYNYLSLTDVGSCTLVC	315
QY	330	-----	329
DB	316	PLHNQVETADGTQRCCKSPCARVCYGLGMEHLREVRVTSANIOEFAGCKIFGSLA	375
QY	330	-----	329
DB	376	FLPESFDGDPASNTAPLOEQVETLEETGYLYISANPDSLPDLVFNQLQVIRGRI	435

```
QY 330 ----- 329
Db 436 LHNGAYSLTLOGLGTSWGLRSLRELGSGLALIHNTLHLCFVHTVPWDQLFRNPHQALLH 495
QY 330 ----- 329
Db 496 TANRPEDECVGEGLACHQLCARGHCWGPGPTQCVNCSQFLRGQECVEBCRVLQGLPREYV 555
QY 330 ----- 329
Db 556 NARHCLPCHPECPQNGSVTCFGEADQCVCAHYKDPFCVARGPSGVKPDLSYMPIWK 615
QY 330 ----- 329
Db 616 FPDEGACQPCPNCTHSCVDLDDKGCAPAEORASPLTSIVSAVVGILLVVLGVVFGILLI 675
QY 330 ----- 329
Db 676 KRRQKIRKYTMRLLOETELVEPLTPSGAMPNQAQMRILKETELRKVKVILGSGAFGTIV 735
QY 330 ----- 329
Db 736 KGIWIPDGENVKIPVAIKVLENTSPKANKELDEAYVMAGVSPVSRLLIGICLTSTVQ 795
QY 330 ----- 329
Db 796 LVTQIMPYGCILLDHVRENRLGSGDQLLNWCQIAKMSYLEYDLVLRHDLAARNVLVKS 855
QY 330 ----- 329
Db 856 PNHVKITDFGLARLLDIDETEHADGGKVPKIKWMALESILRRRTHQSDVMSYGVTVWEL 915
QY 330 ----- 329
Db 916 MTFGAKPYDGIPIAREIPDLLEKGERLPQPPICITDVYMIWKCWIMIDSECRPRFRELVE 975
QY 330 ----- 329
Db 976 FSRWARDPQRFVQIWNEDLGPASPLDSTFYRSLLDDMDGLVDAEEYLVPPQGFPCDP 1035
QY 330 --GAGGMVHRRSSSTRSGGDLTLGLEPSEEBAPRSLAPSGAGSDVFDGLGMGAA 387
Db 1036 APGAGGMVHRRSSSTRSGGDLTLGLEPSEEBAPRSLAPSGAGSDVFDGLGMGAA 1095
QY 388 KGLQSLPHDSPLQRYSEDTVPLPSETDGYVAPLTCSPQPEYVQPDVBPQPSREG 447
Db 1096 KGLQSLPHDSPLQRYSEDTVPLPSETDGYVAPLTCSPQPEYVQPDVBPQPSREG 1155
QY 448 PLPAARPAATLERAKTILSPKNGVVKDVFAGGAVENPEYLTPOGGAAPOPHPPAFSP 507
Db 1156 PLPAARPAATLERAKTILSPKNGVVKDVFAGGAVENPEYLTPOGGAAPOPHPPAFSP 1215
QY 508 AFDNLYWDQDPPPERGAPPSTFKGTPTAENPEYLGLDVP 546
Db 1216 AFDNLYWDQDPPPERGAPPSTFKGTPTAENPEYLGLDVP 1254
```

RESULT 2

```
US-09-811-115-3
; Sequence 3, Application US/09811115
; Patent No. 6632979
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A
; CURRENT APPLICATION NUMBER: US/09/811.115
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 3
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-811-115-3

Query Match      64.5%; Score 2410.5; DB 4; Length 1255;
Best Local Similarity 41.2%; Pred. No. 3.7e-179;
Matches 510; Conservative 0; Mismatches 2; Indels 727; Gaps 1;

QY 35 LARGAATQVCTGDMKRLRPASPEHLDMLRHLYQGCQVVOGNLELTTLPTNASLSFLQ 94
Db 16 LPPGAATQVCTGDMKRLRPASPEHLDMLRHLYQGCQVVOGNLELTTLPTNASLSFLQ 75
QY 95 DIQVQGVVLIHNOVQVPLQRLRIVRGTLQFEDNTALAVLDNGDPLNNTPTPTVGASPG 154
Db 76 DIQVQGVVLIHNOVQVPLQRLRIVRGTLQFEDNTALAVLDNGDPLNNTPTPTVGASPG 135
QY 155 GLRELQRLSLTEILKGGVLIQRPNPQLCYQDTILWKDIFHKNNQALALTIDITNRSRACHPC 214
Db 136 GLRELQRLSLTEILKGGVLIQRPNPQLCYQDTILWKDIFHKNNQALALTIDITNRSRACHPC 195
QY 215 SPMCKGSRGWESSEDCQSLTRTVACGACGACRCKPLPTDCHEQCAAGCTGPKHSDCLAC 274
Db 196 SPMCKGSRGWESSEDCQSLTRTVACGACGACRCKPLPTDCHEQCAAGCTGPKHSDCLAC 255
QY 275 LHFNHSGICELHCPALVTYNTDTPESMPNPEGRYTFGASCVTACPYNYLSTDVGS --- 329
Db 256 LHFNHSGICELHCPALVTYNTDTPESMPNPEGRYTFGASCVTACPYNYLSTDVGSCTLVC 315
QY 330 ----- 329
Db 316 PLHNQVTAEDGTQRCCKPCARVCYGLGMEHLREVRAVTSANIOEAGCKKIFGSLA 375
QY 330 ----- 329
Db 376 FLPEFDDGPASNTAPLQPEQLQVFETLEEITGYLIISAWPDSPLDLSVFQNLQVIRGI 435
QY 330 ----- 329
Db 436 LHNGAYSLTLOGLGTSWGLRSLRELGSGLALIHNTLHLCFVHTVPWDQLFRNPHQALLH 495
QY 330 ----- 329
Db 496 TANRPEDECVGEGLACHQLCARGHCWGPGPTQCVNCSQFLRGQECVEBCRVLQGLPREYV 555
QY 330 ----- 329
Db 556 NARHCLPCHPECPQNGSVTCFGEADQCVCAHYKDPFCVARGPSGVKPDLSYMPIWK 615
QY 330 ----- 329
Db 616 FPDEGACQPCPNCTHSCVDLDDKGCAPAEORASPLTSIVSAVVGILLVVLGVVFGILLI 675
QY 330 ----- 329
Db 676 KRRQKIRKYTMRLLOETELVEPLTPSGAMPNQAQMRILKETELRKVKVILGSGAFGTIV 735
QY 330 ----- 329
Db 736 KGIWIPDGENVKIPVAIKVLENTSPKANKELDEAYVMAGVSPVSRLLIGICLTSTVQ 795
QY 330 ----- 329
Db 796 LVTQIMPYGCILLDHVRENRLGSGDQLLNWCQIAKMSYLEYDLVLRHDLAARNVLVKS 855
QY 330 ----- 329
Db 856 PNHVKITDFGLARLLDIDETEHADGGKVPKIKWMALESILRRRTHQSDVMSYGVTVWEL 915
QY 330 ----- 329
Db 916 MTFGAKPYDGIPIAREIPDLLEKGERLPQPPICITDVYMIWKCWIMIDSECRPRFRELVE 975
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:20:15 ; Search time 75.9936 Seconds
(without alignments)
2565.449 Million cell updates/sec

Title: US-09-821-883-2
Perfect score: 3739
Sequence: 1 MRAAPLLARASLSGLF.....EPVQEGAPPPAAHHHHH 690

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: geneseqp1980s:*

2: geneseqp1990s:*

3: geneseqp2000s:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003as:*

7: geneseqp2003bs:*

8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3739	100.0	690	AAE13109	Human HER
2	3473.5	92.9	697	AAE13111	Human HER
3	2972	79.5	555	AAE13108	Human HER
4	2957.5	79.1	564	AAE13110	Human HER
5	2573.5	68.8	919	AAE21203	Human HER
6	2573.5	68.8	919	AAE51148	HER-2/neu
7	2410.5	64.5	1255	AAE60167	Human HER
8	2410.5	64.5	1255	AAE92620	HER2 tran
9	2410.5	64.5	1255	AAE12130	Human tyr
10	2410.5	64.5	1255	AAE26349	Human HER
11	2410.5	64.5	1255	AAE26366	Human HER
12	2410.5	64.5	1255	AAU74545	Human HER
13	2410.5	64.5	1255	ABR47447	Breast ca
14	2410.5	64.5	1255	ABP74708	Human HER
15	2410.5	64.5	1255	AAE38390	Human c-e
16	2410.5	64.5	1255	ADA38143	Human erb
17	2410.5	64.5	1255	ADA37255	Human erb
18	2410.5	64.5	1255	ADB67621	Human Erb
19	2405.5	64.3	1255	AAW01111	HER-2/neu
20	2405.5	64.3	1255	AAW92406	Human HER
21	2405.5	64.3	1255	AAV84780	Amino aci
22	2405.5	64.3	1255	AAE21198	Human HER
23	2405.5	64.3	1255	AAE88267	HER2/neu
24	2405.5	64.3	1255	AAE85458	Human HER
25	2405.5	64.3	1255	AAE20479	Human HER

26	2405.5	64.3	1255	AAU77114	Human Her
27	2405.5	64.3	1255	AAE51143	Human Her
28	2405.5	64.3	1255	AAE24067	Human Her
29	2405.5	64.3	1255	ABR43687	Human C-e
30	2405.5	64.3	1255	ABR82066	Human Her
31	2405.5	64.3	1255	ADC09593	Her2/Neu
32	2405.5	64.3	1255	ADD25484	Binding d
33	2405.5	64.3	1255	ADG63281	Human Pro
34	2405.5	64.3	1255	ADE76190	Human HER
35	2400.5	64.2	1433	AAE39568	Sequence
36	2398.5	64.1	1253	ADC35106	Human bre
37	2334.5	62.4	1223	AAU98923	Human bre
38	2224	59.5	782	AAW19764	Her2-GM-C
39	2188.5	58.5	479	AAE13112	Human HER
40	2130	57.0	926	AAE51153	Mouse Her
41	2129	56.9	920	AAE51152	Mouse Her
42	2092.5	56.0	1200	AAE21208	Human HER
43	1961	52.4	1256	AAE21206	Mouse Her
44	1961	52.4	1256	AAE62860	Amino aci
45	1961	52.4	1256	AAE51151	Mouse Her

ALIGNMENTS

RESULT 1
AAE13109
ID AAE13109 standard; protein; 690 AA.

XX AAE13109;

AC AAE13109;

DT 28-JAN-2002 (first entry)

DE Human HER500-hgm-CSF fusion protein construct.

KW Immunostimulatory fusion protein; IFP; antigen component; therapy;
KW Immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
KW HER-2 protein; granulocyte-macrophage colony stimulating factor;
KW HER500-hgm-CSF fusion protein.

XX Homo sapiens.

OS Synthetic.

OS Chimeric.

PN WO200174855-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US010515.

XX 30-MAR-2000; 2000US-0193504P.

XX (DEND-) DENDREON CORP.

XX Laus R, Vidovic D, Graddis T;

XX WPI; 2001-662965/76.

XX N-PSDB; AAD21565.

XX An immunostimulatory fusion protein comprising the intracellular domain

XX of HER-2 and an antigen elicits an immune response to the antigen and is

XX useful for the treatment of associated cancer associated.

XX Claim 7; Page 26; 59pp; English.

XX The invention relates to immunostimulatory fusion proteins (IFP) and

XX nucleic acid molecules encoding such proteins. The IFPs comprise a

XX polypeptide antigen component and an immunostimulatory component derived

XX from the intracellular domain of HER-2 protein which is effective to

XX elicit a protective dendritic cell (DC)-induced T-cell mediated cellular

CC immune response to the antigen. IFP or superactivated dendritic cells are
CC used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
CC associated with a particularly antigen. The present sequence is HER500
CC rGM-CSF fusion protein construct which comprises human PAP signal
CC sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal
CC sequence, mature HER-2 membrane distal extracellular and intracellular
CC domains, an Ala Ala linker, a mature human granulocyte- macrophage colony
CC stimulating factor (GM-CSF) sequence and a C-terminal tag
XX
SQ Sequence 690 AA;

Query Match 100.0%; Score 3739; DB 4; Length 690;
Best Local Similarity 100.0%; Pred. No. 7.7e-224;
Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRAAPLLARAAASLSGLFLFLFFWLDRLSVLAKELARGAASQVCTGDMKRLRPASPET 60
DB 1 MRAAPLLARAAASLSGLFLFLFFWLDRLSVLAKELARGAASQVCTGDMKRLRPASPET 60
QY 61 HLDMLRHLVQGCQVQVQGNLELYLPTNASLSFLQDIOEVQGVLIHNVQVPLQRLRI 120
DB 61 HLDMLRHLVQGCQVQVQGNLELYLPTNASLSFLQDIOEVQGVLIHNVQVPLQRLRI 120
QY 121 VRGTLFEDNYALAVLDNGDPLNNTPTVTGASPGGLRELQLRSLEILKGGVLIQRLNPOL 180
DB 121 VRGTLFEDNYALAVLDNGDPLNNTPTVTGASPGGLRELQLRSLEILKGGVLIQRLNPOL 180
QY 181 CYQDTILWKDIFHKNNQALTLIDTNRSRACHPCSPMCKGSRGWESSEDQSLTRTVCA 240
DB 181 CYQDTILWKDIFHKNNQALTLIDTNRSRACHPCSPMCKGSRGWESSEDQSLTRTVCA 240
QY 241 GGCARCKGLPTDCHEQACAGCTGPKHSDCLACLFHNSHIGELHCPALVYNTDTPES 300
DB 241 GGCARCKGLPTDCHEQACAGCTGPKHSDCLACLFHNSHIGELHCPALVYNTDTPES 300
QY 301 MPNPEGRYTFGASCVTACPYNYLSTDVSGAGGVMVHRRHSRSGGDLTLGLEPSEE 360
DB 301 MPNPEGRYTFGASCVTACPYNYLSTDVSGAGGVMVHRRHSRSGGDLTLGLEPSEE 360
QY 361 EAPRSPLAPSGAGSDVFDGLGMAAGKGLQSLPHTDPSPLQRYSEDPTVPLPSETDGYV 420
DB 361 EAPRSPLAPSGAGSDVFDGLGMAAGKGLQSLPHTDPSPLQRYSEDPTVPLPSETDGYV 420
QY 421 APLTCSPOEYVNDVDPQPPSPREGPLPAARPAAGATLERAKTILSPKNGVVDVAFG 480
DB 421 APLTCSPOEYVNDVDPQPPSPREGPLPAARPAAGATLERAKTILSPKNGVVDVAFG 480
QY 481 GAVENPEYLTQGGAAPOPHPPPPAFSPAFDNLVYWDQPPPPGAPPSTFKGTPTAENPEY 540
DB 481 GAVENPEYLTQGGAAPOPHPPPPAFSPAFDNLVYWDQPPPPGAPPSTFKGTPTAENPEY 540
QY 541 LGLDVPAAPARSPPSPQPEHVNAIQEARLLNLSRDTAEMNETVEVISEMFDLOEP 600
DB 541 LGLDVPAAPARSPPSPQPEHVNAIQEARLLNLSRDTAEMNETVEVISEMFDLOEP 600
QY 601 TCIOTRLELYKQGLRGSITLKGPLTMMASHYKQHCPTPTSCATQIITPESKENLKD 660
DB 601 TCIOTRLELYKQGLRGSITLKGPLTMMASHYKQHCPTPTSCATQIITPESKENLKD 660
QY 661 FLLVTFPCWEPVQEGAPPPPAHHHHHHH 690
DB 661 FLLVTFPCWEPVQEGAPPPPAHHHHHHH 690

RESULT 2
AAE13111
ID AAE13111 standard; protein; 697 AA.
XX
AC AAE13111;
XX
DT 28-JAN-2002 (first entry)
XX
DE Human HER500-rGM-CSF fusion construct comprising OVA-derived peptide.

XX Immunostimulatory fusion protein; IFP; antigen component; therapy;
KW immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
KW HER-2 protein; granulocyte-macrophage colony stimulating factor;
KW ovalbumin-derived octapeptide; OVA; rat; HER500-rGM-CSF fusion protein.
XX
OS Homo sapiens.
OS Rattus norvegicus.
OS Unidentified.
OS Chimeric.
PN WO200174855-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US010515.
XX 30-MAR-2000; 2000US-0193504P.
XX (DEND-) DENDREON CORP.
XX Laus R, Vidovic D, Graddis T;
PI WPI; 2001-662965/76.
XX N-PSDB; AAD21567.
XX An immunostimulatory fusion protein comprising the intracellular domain
PT of HER-2 and an antigen elicits an immune response to the antigen and is
PT useful for the treatment of associated cancer associated.
XX
PS Claim 7; Page 27; 59pp; English.
XX
CC The invention relates to immunostimulatory fusion proteins (IFP) and
CC nucleic acid molecules encoding such proteins. The IFPs comprise a
CC polypeptide antigen component and an immunostimulatory component derived
CC from the intracellular domain of HER-2 protein which is effective to
CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
CC immune response to the antigen. IFP or superactivated dendritic cells are
CC used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
CC associated with a particularly antigen. The present sequence is HER500
CC rGM-CSF fusion protein construct which comprises human PAP signal
CC sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal
CC linker, an ovalbumin (OVA)-derived immunodominant octapeptide, an Ala
CC membrane distal intracellular domain, an Ala Ala linker, a mature rat
CC granulocyte-macrophage colony stimulating factor (GM-CSF) sequence and a
CC C-terminal tag
XX
SQ Sequence 697 AA;

Query Match 92.9%; Score 3473.5; DB 4; Length 697;
Best Local Similarity 91.7%; Pred. No. 2.4e-207;
Matches 641; Conservative 19; Mismatches 28; Indels 11; Gaps 2;
QY 1 MRAAPLLARAAASLSGLFLFLFFWLDRLSVLAKELARGAASQVCTGDMKRLRPASPET 60
DB 1 MRAAPLLARAAASLSGLFLFLFFWLDRLSVLAKELARGAASQVCTGDMKRLRPASPET 60
QY 61 HLDMLRHLVQGCQVQVQGNLELYLPTNASLSFLQDIOEVQGVLIHNVQVPLQRLRI 120
DB 61 HLDMLRHLVQGCQVQVQGNLELYLPTNASLSFLQDIOEVQGVLIHNVQVPLQRLRI 120
QY 121 VRGTLFEDNYALAVLDNGDPLNNTPTVTGASPGGLRELQLRSLEILKGGVLIQRLNPOL 180
DB 121 VRGTLFEDNYALAVLDNGDPLNNTPTVTGASPGGLRELQLRSLEILKGGVLIQRLNPOL 180
QY 181 CYQDTILWKDIFHKNNQALTLIDTNRSRACHPCSPMCKGSRGWESSEDQSLTRTVCA 240
DB 181 CYQDTILWKDIFHKNNQALTLIDTNRSRACHPCSPMCKGSRGWESSEDQSLTRTVCA 240

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:57:41 ; Search time 3742.37 Seconds
(without alignments)
4428.608 Million cell updates/sec

Title: US-09-821-883-1
Perfect score: 3020
Sequence: 1 MRAAPLLARAASLSGLFLF.....ENPEYGLDVPAAAHHHH 555

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09821883/runat_09092004_105127_6643/app_query.fasta_1.3100
-DB=EST -QFMT=fastcap -SUFFIX=rc -MINMATCH=0.1 -LOOPEXT=0 -TRANS=human40.cdi -List=45
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -MINLEN=15 -MODE=LOCAL
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09821883 @CGN 1 1 11571 @runat_09092004_105127_6643 -NCPU=6 -ICPU=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=110 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1964	65.0	4323	11	AK031099	AK031099 Mus muscu
2	1947.5	64.5	4463	11	AK083669	AK083669 Mus muscu
3	1288	42.6	1016	12	BM562913	BM562913 AGENCOURT
4	1175	38.9	1129	12	BM802792	BM802792 AGENCOURT
5	1163	38.5	871	14	CA488274	CA488274 AGENCOURT
6	1163	38.5	885	14	CA455074	CA455074 AGENCOURT
7	1156	38.3	4715	11	AF318349	AF318349 Homo sapi
8	1137.5	37.7	893	14	CA455141	CA455141 AGENCOURT
9	1090	36.1	1004	14	CD515356	CD515356 AGENCOURT
10	1087	36.0	894	14	CA454131	CA454131 AGENCOURT
11	1080	35.8	653	12	BM721340	BM721340 UI-E-E01-
12	1073	35.5	588	13	BX478931	BX478931 DKFZp686A
13	1072	35.5	902	14	CA488868	CA488868 AGENCOURT
14	1046	34.6	569	9	AL701765	AL701765 DKFZp686P
15	1027	34.0	583	13	BX479114	BX479114 DKFZp686P
16	1016.5	33.7	795	14	CA328613	CA328613 UI-M-PY0-
17	1012.5	33.5	3110	11	AK031542	AK031542 Mus muscu
18	1004	33.2	808	14	CA489534	CA489534 AGENCOURT
19	997	33.0	616	14	CD721801	CD721801 oJ1f07.Y
20	976	32.3	943	13	BQ958632	BQ958632 AGENCOURT
21	976	32.3	1988	11	BC023725	BC023725 Mus muscu
22	976	32.3	3372	11	BC046553	BC046553 Mus muscu
23	931	30.8	813	12	BI557797	BI557797 603236977
24	926	30.7	855	13	BU594980	BU594980 AGENCOURT
25	915.5	30.3	852	14	CD562283	CD562283 AGENCOURT
26	915	30.3	614	10	AW370693	AW370693 QV1-BT026
27	912	30.2	791	12	BI154872	BI154872 602902857
28	910	30.1	717	14	CB598701	CB598701 AGENCOURT
29	901	29.8	491	12	BM790293	BM790293 K-EST0070
30	886	29.3	906	14	CA454570	CA454570 AGENCOURT
31	883	29.2	998	12	BI649877	BI649877 603296516
32	881	29.2	685	9	AU123871	AU123871 AU123871
33	879.5	29.1	932	14	CA487981	CA487981 AGENCOURT
34	868	28.7	609	10	AW701942	AW701942 uq93b02.Y
35	862	28.5	621	12	BG283493	BG283493 602407782
36	855	28.3	786	12	BI155788	BI155788 602904360
37	846	28.0	887	13	BQ717097	BQ717097 AGENCOURT
38	835.5	27.7	649	9	AA496412	AA496412 zv37a02.f
39	817	27.1	533	14	CB437620	CB437620 685491.MA
40	817	27.1	552	12	BG277542	BG277542 ux45d07.Y
41	811	26.9	670	9	AI906012	AI906012 RC-BT105-
42	808	26.8	876	13	BQ769889	BQ769889 UI-M-P10-
43	803	26.6	887	14	CA980253	CA980253 AGENCOURT
44	802	26.6	514	10	BF998814	BF998814 QV0-GN014
45	797.5	26.4	678	12	BI555157	BI555157 603236396

ALIGNMENTS

RESULT 1

AK031099

LOCUS

DEFINITION

AK031099 Mus musculus 13 days embryo forelimb cDNA, RIKEN full-length enriched library, clone:530404N10 product:v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian), full insert sequence.

ACCESSION

AK031099.1 GI:26082143

VERSION

AK031099.1 HT:26082143

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayaishizaki, Y.
High-efficiency full-length cDNA cloning
JOURNAL 99279253
MEDLINE 10349636
PUBMED

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayaishizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL 20499374
MEDLINE 11042159
PUBMED

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, I., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayaishizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL 20530913
MEDLINE 11076861
PUBMED

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 585-590 (2001)
JOURNAL

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
JOURNAL 20530913
MEDLINE 11076861
PUBMED

6 (bases 1 to 4323)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayaishida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayaishizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayaishizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.
Location/Qualifiers
1. .4323
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

/db_xref="FANTOM DB:5930404N10"
/db_xref="MGI:2394277"
/db_xref="taxon:10090"
/clone="5930404N10"
/tissue_type="forelimb"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
48. .3818
misc_feature
note="putative
v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian)
(MGI:MGI:95410, GB|U71126, evidence: BLASTN, 99%, match=449)"

ORIGIN

Alignment Scores:
Pred. No.: 1.25e-112 Length: 4323
Score: 1964.00 Matches: 437
Percent Similarity: 36.17% Conservative: 26
Best Local Similarity: 34.14% Mismatches: 77
Query Match: 65.03% Indels: 740
DB: 11 Gaps: 3

US-09-821-883-1 (1-555) x AK031099 (1-4323)

Qy 3 AlaAlaProLeuLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPheLeuLeu 22
Db 33 GCCGACGATATCATCATGGAGCTGGCGCTGGTGGGGGTTCTCTCGCCCTC 92
Qy 23 PhePheTrpLeuAepArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAlaSerThr 42
Db 93 -----CTGTCCCCCGAGCGCGGTACC 116
Qy 43 GlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThrHisLeu 62
Db 117 CAAAGTGTGTACCGGTACCGATGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 176
Qy 63 AspMetLeuArgHisLeuTyrglnGlyCysGlnValValGlnGlnGlnGlnGlnGlnGln 82
Db 177 GACATGTTTGGCCACCTCTACCGAGGCTGTTCAGGTGTGTCAGGGCAATTTGGAGCTTACC 236
Qy 83 TyrlleuProThrAenAlaSerLeuSerPheLeuGlnAspIleGlnGlnGlnGlnGlnGln 102
Db 237 TACCTGCGCCCAATGCGAGCTCTCATCTCTGTCAGGACATCCAGGAAGTCCAGGGATAC 296
Qy 103 ValIleuIleAlaHisAenGlnValArgGlnValProLeuGlnArgLeuArgGlnValArg 122
Db 297 ATGCTCATCTGCTCACAACCGAGTGAACACGCTCCCACTGTCAGAGGTTGCGCATCGTGAGA 356
Qy 123 GlyThrGlnLeuPheGluAspAsnTyrglnAlaLeuAlaValLeuAspAsnGlyAspProLeu 142
Db 357 GGGACTGAGCTCTTTGAGGACAGTATGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 416
Qy 143 AsnAen---ThrThrProValThrGlyAlaSerProGlyGlyLeuArgGlnGlnLeu 161
Db 417 GACAACTGACACCGCGCCCGCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 476
Qy 162 ArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeuCys 181
Db 477 CGAAGTCTCAGAGATCTTTGAAGGGAGGAGTGTGTGATGCTGGGAACCCCTCAGCTCTGC 536
Qy 182 TyrglnAspThrIleLeuTyrLysAspIlePheHisLysAsnAenGlnLeuAlaLeuThr 201
Db 537 TACGAGGACATGTTTGTGAGAGGATGCTCCGTAAGATATACAGAGTGGCTCTGCTGTC 596
Qy 202 LeuIleAspThrAenArgSerArgAlaCysHisProCysSerProMetCysLysGlySer 221
Db 597 GACATGGACACCAATCGTTCGGGGCTGTCCACCTTGTGTGCCCCCAACCTGCAAGACAAT 656
Qy 222 ArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAlaGly 241
Db 657 CACTGTTGGGGGTGAGAGTCTCTGAAGACTGTGAGATCTTACTTGGCACCCTCTGACTAGT 716

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 15:10:16 ; Search time 671.049 Seconds
(without alignments)
4161.385 Million cell updates/sec

Title: US-09-821-883-1

Perfect score: 3020

Sequence: 1 MRAAPLLARASLSGLF.....ENPEYGLDVPAAHHHH 555

Scoring table:

	BLOSUM62	Xgapop 10.0	Xgapext 0.5
Xgapop 10.0			
Xgapext 0.5			
Fgapop 6.0			
Delep 6.0			

Searched: 3304383 seqs, 2515761380 residues

Total number of hits satisfying chosen parameters: 6608766

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09821883/runat_09092004_105129_6736/app_query.fasta_1.3100
-DB=Published Applications_NA_QFMT=fastap -SUFFIX=znpp -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09821883 @CGN 1.1949 @runat_09092004_105129_6736
-NCPU=6 -ICPU=3 -NO_MMAP -LARGESQUERY -NEG SCORES=0 -WAIT_DSBELOC=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRAD=10 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
------------	-------	--------------	----	-------------

ALIGNMENTS

RESULT 1

US-09-821-883-6
; Sequence 6, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821.883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500 construct

	1	3020	100.0	1665	9	US-09-821-883-6	Sequence 8, Appl
1	3020.5	99.5	1692	9	US-09-821-883-8	Sequence 6, Appl	
2	3005.5	99.5	1692	9	US-09-821-883-8	Sequence 8, Appl	
3	2972	98.4	2070	9	US-09-821-883-7	Sequence 7, Appl	
4	2957.5	97.9	2091	9	US-09-821-883-9	Sequence 9, Appl	
5	2419	80.1	4543	9	US-09-759-508-1	Sequence 1, Appl	
6	2416	80.0	4530	9	US-09-877-177-11	Sequence 11, Appl	
7	2416	80.0	4530	13	US-10-392-113-45	Sequence 45, Appl	
8	2416	80.0	4530	13	US-10-426-836-11	Sequence 11, Appl	
9	2416	80.0	4530	13	US-10-177-293-125	Sequence 125, Appl	
10	2416	80.0	4530	15	US-10-007-926A-119	Sequence 119, Appl	
11	2416	80.0	4530	15	US-10-338-730-1	Sequence 1, Appl	
12	2416	80.0	4530	15	US-10-101-510-124	Sequence 124, Appl	
13	2416	80.0	4530	16	US-10-116-275-131	Sequence 131, Appl	
14	2416	80.0	4530	16	US-10-272-437A-27	Sequence 27, Appl	
15	2416	80.0	4530	16	US-10-117-937-595	Sequence 595, Appl	
16	2416	80.0	4530	16	US-10-159-563-208	Sequence 208, Appl	
17	2416	80.0	4530	16	US-10-435-696-10	Sequence 10, Appl	
18	2416	80.0	4530	17	US-10-734-564-59	Sequence 59, Appl	
19	2416	80.0	4642	15	US-10-198-846-10896	Sequence 10896, A	
20	2413.5	79.9	9274	9	US-09-811-123-7	Sequence 7, Appl	
21	2413.5	79.9	9274	9	US-09-811-115-1	Sequence 1, Appl	
22	2411	79.8	4473	10	US-09-441-411-5	Sequence 5, Appl	
23	2411	79.8	4473	15	US-10-146-473-32	Sequence 32, Appl	
24	2411	79.8	4473	15	US-10-207-655-44	Sequence 44, Appl	
25	2411	79.8	4473	15	US-10-101-510-81	Sequence 81, Appl	
26	2410.5	79.8	3765	15	US-10-207-498-5	Sequence 5, Appl	
27	2410.5	79.8	3768	9	US-09-811-123-8	Sequence 8, Appl	
28	2410.5	79.8	3768	9	US-09-811-115-2	Sequence 2, Appl	
29	2410.5	79.8	3768	13	US-09-984-092-3	Sequence 3, Appl	
30	2410.5	79.8	3768	13	US-10-280-576-3	Sequence 3, Appl	
31	2410.5	79.8	3768	17	US-10-441-779C-3	Sequence 3, Appl	
32	2405.5	79.7	3768	9	US-09-854-356-9	Sequence 9, Appl	
33	2405.5	79.7	3768	9	US-09-930-125-1	Sequence 1, Appl	
34	2405.5	79.7	3768	13	US-09-765-973-1	Sequence 1, Appl	
35	2405.5	79.7	3768	15	US-10-313-644-1	Sequence 1, Appl	
36	2386	79.0	4606	10	US-09-971-392-70	Sequence 70, Appl	
37	1961	64.9	3771	9	US-09-854-356-11	Sequence 11, Appl	
38	1956.5	64.8	3955	9	US-09-870-759-117	Sequence 117, Appl	
39	1956.5	64.8	3955	9	US-09-854-356-10	Sequence 10, Appl	
40	1956.5	64.8	3955	10	US-09-751-708A-117	Sequence 117, Appl	
41	1795	59.4	1437	9	US-09-821-883-10	Sequence 10, Appl	
42	1641.5	54.4	2132	16	US-10-412-804A-3	Sequence 3, Appl	
43	1639	54.3	2149	16	US-10-412-804A-9	Sequence 9, Appl	
44	1636	54.2	2164	16	US-10-412-804A-5	Sequence 5, Appl	
45	1587	52.5	867	9	US-09-821-883-24	Sequence 24, Appl	

US-09-821-883-6

Alignment Scores:

Pred. No.: 4,93e-272 Length: 1665
Score: 3020.00 Matches: 555
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-821-883-1 (1-555) x US-09-821-883-6 (1-1665)

QY 1 MetArgAlaAlaProLeuLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPhe 20
DB 1 ATGAGAGCTGACCCCTCTCTCTGCGCAGGCGAGCAGCTTAGCCTTGGCTTCTGTTT 60
QY 21 LeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAla 40
DB 61 CTGCTTTTCTGGCTAGACCGAAGTGTACTAGCCAAAGGAGTTGGCGCGCGCGCG 120
QY 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
DB 121 TCAGCCCAAGTGTGACCGGCACAGATGAAGCTGCGGCTCCCTGCGAGTCCCGAGACC 180
QY 61 HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnAsnLeuGlu 80
DB 181 CACTTGGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGTGCGAGGANAACCTGGAA 240
QY 81 LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGlnValGln 100
DB 241 CTCACCTACCTGCCCAACCAATGCCAGCTGTCTCTCTCCAGGATATCCAGGAGTGCAG 300
QY 101 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
DB 301 GGCTACGTGCTCATCGCTCAACCAAGTGGGAGTCCACCTGCGAGAGGCTGCGGATT 360
QY 121 ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 140
DB 361 GTGGAGGACCCAGCTCTTTGAGGACACTATGCCCTGCGCTGCTAGACATGGAGAC 420
QY 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln 160
DB 421 CCGCTGAACAATACCAACCCCTGTCAAGGGGCTCCCGAGGGCTGCGGAGCTGCAG 480
QY 161 LeuArgSerLeuThrGluLeuLeuLysGlyValLeuIleGlnArgAsnProGlnLeu 180
DB 481 CTTGGAAGCTTCACAGAGATCTTGAAAGGAGGGGTCTTGATCCAGCGAAACCCCGAGCTC 540
QY 181 CysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200
DB 541 TGCTACGAGGACAGATTTTGTGAAGAGACATCTTCCACAGAAACAACAGCTGGCTCTC 600
QY 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220
DB 601 ACCTGTAGACACCAACCGCTCTCGGCGCTGCCACCCCTGTTCTCCGATGTGTAAGGCG 660
QY 221 SerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAla 240
DB 661 TCCCGCTGCTGGGAGAGAGTTCTGAGGATTTGTGAAGAGACATCTTCCACAGAAACAACAGCTGGCTCTC 720
QY 241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCysAla 260
DB 721 GGTGGCTGTGCCCTGCAAGGGGCCACTGCGCCACTGACTGCTGCTGCTGCTGCTGCT 780
QY 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
DB 781 GCGGCTGCAGCGGCCCAAGCACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
QY 281 GlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSer 300
DB 841 GGCATCTGTGAGCTGCACTGCCAGCCCTGGTCACTACACACAGACAGACGTTTGAGTCC 900
QY 301 MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr 320

DB 901 ATGCCCAATCCCGAGGCGCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGTCCCTAC 960
QY 321 AsnTyrLeuSerThrAspValGlySerGlyAlaGlyGlyMetValHisHisArgHisArg 340
DB 961 AACTACTCTTTCTACCGACGTGGGATCGGGCGCTGGGGGATGGTCCACACAGGACCGC 1020
QY 341 SerSerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGlu 360
DB 1021 AGCTCATCTACAGAGGTGGCGGTGGGACCTGACACTAGGGGTGGAGCCCTCTGAAGAG 1080
QY 361 GluAlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGly 380
DB 1081 GAGGCCCCAGGTCTCCACTGGCACCTCCGAAGGGGCTGGCTCCGATGTATTGATGGT 1140
QY 381 AspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerPro 400
DB 1141 GACCTGGGAATGGGGGAGCAAGGGGCTGCAAGGCTCCCCACACATGATGCCCGCCT 1200
QY 401 LeuGlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrVal 420
DB 1201 CTACAGGGGTACAGTGAGGACCCCACTACCTCCCTGCGCTCTGAGACTGATGGCTACGTT 1260
QY 421 AlaProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGln 440
DB 1261 GCCCCCTGACCTGCAGCCCGCCAGCCTGAATATGTGAACCGACAGATGTTGGGCCCGAG 1320
QY 441 ProProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlu 460
DB 1321 CCCCTTTCGGCCCGAGAGGGCCCTCTGCGCTGCGCCGACCTGCTGTGCTGCTCTGGAA 1380
QY 461 ArgAlaLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly 480
DB 1381 AGGCCCAAGACTCTCTCCCGAGGAGAGATGGGGTGTCAAAGACGTTTTTTCCTTTGGG 1440
QY 481 GlyAlaValGluAsnProGluTyrLeuThrProGlnGlyGlyAlaAlaProGlnProHis 500
DB 1441 GGTGCGGTGGAGAACCCGAGTACTTGACACCCCGAGGAGGAGCTGCGCTCAGCCCCAC 1500
QY 501 ProProAlaPheSerProAlaPheAspAsnLeuTyrTyrTrpAspGlnAspProPro 520
DB 1501 CCTCTCTCTGCTTCAGCCCGAGCCTTCGACCACTTATTACTGGGACAGGACCCACCA 1560
QY 521 GluArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyr 540
DB 1561 GAGCGGGGGCTCCACCCAGCCTTCAAAGGGACACTACGGCAGAGAACCCAGAGTAC 1620
QY 541 LeuGlyLeuAspValProAlaAlaAlaHisHisHisHisHisHisHisHisHis 555
DB 1621 CTGGGTCTGACGCTGCCGCGCGCGACATCACCATCACCATCAC 1665

RESULT 2

US-09-821-883-8
; Sequence 8, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; TITLE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1692
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 06:51:36 ; Search time 95.6744 Seconds
(without alignments)
3219.229 Million cell updates/sec

Title: US-09-821-883-1
Perfect score: 3020
Sequence: 1 MRAPILLARASLSLGLFLP.....ENPEYLGLDVPAAAHHHHH 555

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=x1h
-Q/cgn2_1/USFTO_spool/US09821883/runat_09092004_105128_6657/app_query.fasta_1.3100
-DB=issued Patents NA -Qfmt=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCLD=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USFR=US09821883 @CGN 1 1 215 @runat_09092004_105128_6657 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Databases : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2416	80.0	4530	1	US-08-229-515A-9
2	2416	80.0	4530	1	Sequence 9, Appli
3	2416	80.0	4530	4	US-08-645-865-9
4	2416	80.0	4530	4	Sequence 4, Appli
5	2416	80.0	4530	4	US-03-527-487-1
6	2413.5	79.9	9274	4	US-08-877-177A-11
7	2411	79.8	4473	2	US-09-811-115-1
8	2411	79.8	4473	3	US-09-056-105-26
9	2411	79.8	4473	4	US-09-663-834A-3
10	2410.5	79.8	3768	4	US-08-811-115-2
11	2405.5	79.7	3768	2	US-08-625-101-1
12	2405.5	79.7	3768	2	US-08-356-786-1

13	1956.5	64.8	3955	1	US-08-229-515A-14
14	1956.5	64.8	3955	1	US-08-645-865-14
15	1639	54.3	2385	2	US-09-146-283-3
16	1639	54.3	2385	3	US-08-579-823A-3
17	1639	54.3	2385	3	US-09-344-195-3
18	1624	53.8	1872	3	US-08-422-108-2
19	1624	53.8	1872	4	US-08-422-734-2
20	725.5	24.0	1868	1	US-08-658-883B-1
21	725.5	24.0	1868	4	US-09-676-610B-26
22	724.5	24.0	1593	4	US-09-676-610B-25
23	724.5	24.0	5532	2	US-08-475-035-3
24	724.5	24.0	5532	2	US-09-676-610B-17
25	721	23.9	5484	3	US-09-632-580A-3
26	721	23.9	5501	1	US-08-484-438-1
27	721	23.9	5555	1	US-08-484-438-3
28	714.5	23.7	4905	1	US-07-978-895-3
29	714.5	23.7	4905	1	US-08-473-119-3
30	714.5	23.7	4905	2	US-08-475-352-3
31	714.5	23.7	4905	4	US-09-170-699-3
32	714.5	23.7	4975	3	US-09-630-706-3
33	705.5	23.4	1958	4	US-09-570-454-1
34	705.5	23.4	1958	4	US-09-867-521-1
35	623	20.6	4545	6	5183884-3
36	493	16.3	322	1	US-08-421-356-1
37	493	16.3	322	4	US-09-046-783-1
38	260.5	8.6	4149	2	US-08-737-715-1
39	260.5	8.6	4723	4	US-09-023-655-1137
40	257	8.5	2533	3	US-09-149-922-5
41	252.5	8.4	4975	2	US-08-249-687C-1
42	252.5	8.4	4989	2	US-08-666-392A-3
43	252.5	8.4	4989	2	US-08-625-819-1
44	252.5	8.4	4989	3	US-08-755-558-4
45	252.5	8.4	4989	3	US-08-746-559A-1

ALIGNMENTS

RESULT 1

US-08-229-515A-9
; Sequence 9, Application US/08229515A
; Patent No. 5518885
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,515A
; FILING DATE: 19 APR 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs

;
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-229-515A-9

Alignment Scores:

Pred. No.: 1,67e-178 Length: 4530
Score: 2416.00 Matches: 520
Percent Similarity: 41.14% Conservative: 5
Best Local Similarity: 40.75% Mismatches: 17
Query Match: 80.00% Indels: 734
DB: 1 Gaps: 3

US-09-821-883-1 (1-555) x US-08-229-515A-9 (1-4530)

Qy 2 ArgAlaAlaProLeuLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17
Db 94 CGGCGCCCTCCAGCGGGTCCAGCCGAGCATGGGGCCGAGCGCAGTGAGCACCATG 153
Qy 18 PheLeuPheLeuLeuPhePheTyrLeuAspArgSerValLeuAlaLysGluLeuAlaArg 37
Db 154 GAGTGGCGGCTTGTGCCGCTG-----GGGCTCCTCTCGGCCCTTTCGCCCCC 204
Qy 38 GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57
Db 205 GGAGCCGCGAGCACCAGTGTGCACCGGCACACATGAGCTGGGCTCCCTGCCAGT 264
Qy 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGly 77
Db 265 CCGGAGACCCACCTGGACATGCTCCGCCACCTCTTACCAGGGCTGCCAGGTGGTGCAGGA 324
Qy 78 AsnLeuGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97
Db 325 AACCTGGAATCACTACCTACCTGCCCCACCAATGCCAGCCTGCTTCTTGCAGGATATCCAG 384
Qy 98 GluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg 117
Db 385 GAGTGCAGGGCTACGTGCTCATCGCTCACACCAAGTAGGCGAGGTCCCACCTGCAGAGG 444
Qy 118 LeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAsp 137
Db 445 CTGCGGATTGTGGAGGACCCAGCTCTTTGAGGACAACTATGCGCTTGGCGGTGTAGAC 504
Qy 138 AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157
Db 505 AATGGAGACCCGCTGAACAATACACCCCTGTACAGGGCCCTCCCGAGGAGGCTTCGG 564
Qy 158 GluLeuGlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsn 177
Db 565 GAGCTGCAGCTTCGAGCCTCCACAGATCTTGAAGAGGGGTCTTGATCCAGCGGAAC 624
Qy 178 ProGlnLeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGln 197
Db 625 CCCAGCTCTGCTACCAAGGACACGATTTGTGGAAGGACATCTTCCACAAGAAACAACAG 684
Qy 198 LeuAlaLeuThrIleLeuAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217
Db 685 CTGGCTCTCACTGATAGACACCAACCGCTCTCGGGCCCTGCCACCCCTGTTCGGATG 744
Qy 218 CysLysGlySerArgCysTyrGlyGluSerSerGluAspCysGlnSerLeuThrArgThr 237
Db 745 TGTAAAGGGCTCCCGCTGCTGGGGAGAGATTTGTGAGGATTTGTACAGCCTGACCGGACT 804
Qy 238 ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlu 257
Db 805 GTCTGTCCGGTGGCTGTGCCGCTGCAGAGGGGCCACTGCCCACTGACTGCTGCCATGAG 864
Qy 258 GlnCysAlaIaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277
Db 865 CAGTGTGCTCCGGCTGCACGGGGCCCAAGACCTCTGACTGCTGCGCTGCCCTCCACTTC 924
Qy 278 AsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThr 297

Db 925 AACCAACAGTGGCATCTGTGAGCTGCACATGCCACGCCCTGTGTACACACACAGACACG 984
Qy 298 PheGluSerMetProAsnProGluGlyArgTyrThrPheGlyValaSerCysValThrAla 317
Db 985 TTTGAGTCCATGCCCCAATCCGAGGGCCGGTATACATTCTGGGCCAGCTGTGTGACTGCC 1044
Qy 318 CysProTyrAsnTyrLeuSerThrAspValGlySer----- 329
Db 1045 TGTCCCTCAACAACCTACCTTTCTACGGACGTGGATCTCTGCACCCCTCGTCTGCCCTGCAC 1104
Qy 329 ----- 329
Db 1105 AACCAAGAGGTGACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGCAAGCCCTGT 1164
Qy 329 ----- 329
Db 1165 GCCGAGTGTGCTATGGTCTGGGCATGGAGCATCTTGGAGAGGTGAGGGCAGTTACCACT 1224
Qy 329 ----- 329
Db 1225 GCCAATATCCAGGAGTTTGTGCTGCAGGAAGATCTTTGGGAGCCTTGGCATTTCTGCCG 1284
Qy 329 ----- 329
Db 1285 GAGAGCTTTGATGGGACCCAGCCTCCACACACTGCCCGCTCCAGCCAGCAGAGCTCCAA 1344
Qy 329 ----- 329
Db 1345 GTTTTGGAGACTCTGGAAGAGATCAAGGTTACTATATATCTCAGCATGCGCGGACAGC 1404
Qy 329 ----- 329
Db 1405 CTGCTGACCTCAGCGCTCTTCCAGAACCTGCAAGTAATCCGGGACGAATTTCTGCACAAT 1464
Qy 329 ----- 329
Db 1465 GGCGCCTACTCTGCTGACCCCTGCAAGGGCTGGGCATCAGCTGGCTGGGGCTCGCGTCACTG 1524
Qy 329 ----- 329
Db 1525 AGGGAATCGGGCAGTGGATCGGCCCTCATCCACATAACACCACTCTGCTTCTGCTGCAC 1584
Qy 329 ----- 329
Db 1585 ACGGTGCCCTGGGACACAGCTCTTTTCGGAACCCGACCAAGCTCTGTCTCCACACTGCCAAC 1644
Qy 329 ----- 329
Db 1645 CGGCCAGAGGACGAGTGTGTGGCGAGGGCTTGGCCCTGCCACCAAGCTGTGCGCCCGAGGG 1704
Qy 329 ----- 329
Db 1705 CACTGTGGGGTCCAGGGGCCACCCAGTGTGTCACTGCAGCAGATTCCTTCGGGGGCCAG 1764
Qy 329 ----- 329
Db 1765 GAGTGTGGAGGAATCCCGAGTACTGCAGGGGCTCCCGAGGAGTATGTGAATGCCAGG 1824
Qy 329 ----- 329
Db 1825 CACTGTTTGGCGTGCCACCCCTGAGTGTGAGCCCCAGAAATGGCTCAGTGACCTGTTTTTGA 1884
Qy 329 ----- 329
Db 1885 CCGGAGGCTGACACAGTGTGTGGCTGTGCCCCACTATAAGGACCCCTCCTTCTGCGTGGCC 1944
Qy 329 ----- 329
Db 1945 CGCTGCCCCAGCGGTGTGAAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGAT 2004
Qy 329 ----- 329
Db 2005 GAGGAGGGGCGATGCCAGCCTTGGCCCCATCAACTGCACCCCACTCTGTGTGGACCTGGAT 2064

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:04:16 ; Search time 546.584 Seconds
(without alignments)
4313.608 Million cell updates/sec

Title: US-09-821-883-1

Perfect score: 3020

Sequence: 1 MRAPLLARASLSGLF.....ENPEYGLDVPAHHHHH 555

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09821883/runat.09092004.105126.6623/app_query.fasta_1.3100
-DB=N_Geneseq_29Jan04 -QMT=fastacp -SUFFIX=ring -MINMATCH=0.1 -LOOFCU=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09821883.qcgn.1.1646@runat.09092004.105126.6623 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=1120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2003as:*
8: Geneseq2003bs:*
9: Geneseq2003cs:*
10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3020	100.0	1665	4 AAD21564	Aad21564 Human HER
2	3005.5	99.5	1692	4 AAD21566	Aad21566 Human HER
3	2972	98.4	2070	4 AAD21565	Aad21565 Human HER
4	2957.5	97.9	2091	4 AAD21567	Aad21567 Human HER
5	2416	80.0	4530	2 AAT01585	Aat01585 Her-2/neu
6	2416	80.0	4530	2 AAT71253	Aat71253 Human HER
7	2416	80.0	4530	3 AAZ60815	Aaz60815 Nucleotid
8	2416	80.0	4530	4 AAD19731	Aad19731 Human tyr

9	2416	80.0	4530	6 ABN85585	Abn85585 Human HER
10	2416	80.0	4530	6 ABZ35012	Abz35012 Human gen
11	2416	80.0	4530	6 ABV94128	Abv94128 Breast ca
12	2416	80.0	4530	6 ABK83918	Abk83918 Human cDN
13	2416	80.0	4530	7 ACC50139	Acc50139 Breast ca
14	2416	80.0	4530	7 ABQ83856	Abq83856 Human Her
15	2416	80.0	4530	8 AAD58073	Aad58073 Human c-e
16	2416	80.0	4530	9 ADC09594	Adc09594 Her2/Neu
17	2413.5	79.9	9274	6 RAD43934	Rad43934 HER-2 tra
18	2413.5	79.9	9274	6 ABK14057	Abk14057 Human HER
19	2411	79.8	4472	3 AAA14812	Aaa14812 cDNA enco
20	2411	79.8	4473	2 ABQ76220	Abq76220 Human tum
21	2411	79.8	4473	2 AAZ31071	Aaz31071 HER-2 nuc
22	2411	79.8	4473	6 ABZ34969	Abz34969 Human gen
23	2411	79.8	4473	6 RAD38904	Rad38904 Human Her
24	2411	79.8	4473	7 ACC69999	Acc69999 Human c-e
25	2411	79.8	4473	9 ADC35148	Adc35148 Human bre
26	2411	79.8	4473	9 ADD25483	Add25483 Binding d
27	2410.5	79.8	3765	9 ADB67620	Adb67620 Human epl
28	2410.5	79.8	3768	3 AAA09455	Aaa09455 Human her
29	2410.5	79.8	3768	6 ABV78168	Abv78168 Human ERB
30	2410.5	79.8	3768	6 ABZ35744	Abz35744 Human ERB
31	2410.5	79.8	3768	6 RAD43935	Rad43935 Human HER
32	2410.5	79.8	3768	6 ABX09987	Abx09987 Human ERB
33	2410.5	79.8	3768	6 AAD43986	Aad43986 Human Her
34	2410.5	79.8	3768	6 ABK14058	Abk14058 Human HER
35	2410.5	79.8	3768	6 ABL91709	Ab191709 Human pol
36	2410.5	79.8	3768	7 ACC57649	Acc57649 Human pro
37	2406.5	79.7	9274	4 AAF24297	Aaf24297 HER2 tran
38	2405.5	79.7	3768	2 AAT40739	Aat40739 HER-2/neu
39	2405.5	79.7	3768	2 AAX01912	Aax01912 Human HER
40	2405.5	79.7	3768	5 AAH23392	Aah23392 Human HER
41	2405.5	79.7	3768	6 RAD32743	Rad32743 Human Her
42	2405.5	79.7	3768	6 ABK10730	Abk10730 Human Her
43	2405.5	79.7	3768	6 ABA92250	Abag2250 Human Her
44	2389.5	79.1	4299	2 ABA46083	Aba46083 Sequence
45	2386	79.0	4606	9 ADB47370	Adb47370 Human CDN

ALIGNMENTS

RESULT 1
AAD21564
ID AAD21564 standard; DNA; 1665 BP.

XX AC AAD21564;

XX DT 28-JAN-2002 (first entry)

XX DE Human HER500 fusion DNA construct.

XX KW Immunostimulatory fusion protein; IFP; antigen component; therapy;
KW Immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human;
KW HER-2 protein; HER500 fusion DNA; ds.

XX OS Homo sapiens.

XX OS Synthetic.

XX OS Chimeric.

XX FH Key
XX CDS Location/Qualifiers
1..1665
/tag=a

FT /product= "Human HER500 fusion protein construct
FT comprising human PAP signal sequence, mature PAP protein,
FT an Ala Arg linker, human HER-2 signal sequence, mature
FT HER-2 membrane distal extracellular and intracellular
FT domains and a C-terminal tag"
FT /note= "CDS does not include stop codon"
FT /partial
XX

```

PN WO200174855-A2.
XX
XX PD 11-OCT-2001.
XX
XX PF 30-MAR-2001; 2001WO-US010515.
XX
XX PR 30-MAR-2000; 2000US-0193504P.
XX
XX PA (DEND-) DENDREON CORP.
XX
XX PI Laus R, Vidovic D, Graddis T;
XX
XX DR WPI; 2001-662965/76.
XX
XX DR P-PSDB; AAE13108.
XX
XX PT An immunostimulatory fusion protein comprising the intracellular domain
XX of HER-2 and an antigen elicits an immune response to the antigen and is
XX useful for the treatment of associated cancer associated.
XX
XX PS Example 3; Page 27; 59pp; English.
XX
XX CC The invention relates to immunostimulatory fusion proteins (IPF) and
XX nucleic acid molecules encoding such proteins. The IPFs comprise a
XX polypeptide antigen component and an immunostimulatory component derived
XX from the intracellular domain of HER-2 protein which is effective to
XX elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
XX immune response to the antigen. IPF or superactivated dendritic cells are
XX used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
XX associated with a particularly antigen. The present sequence is HER500
XX fusion DNA construct which comprises DNA molecules encoding human PAP
XX signal sequence, mature PAP protein, an Ala Arg linker, human HER-2
XX signal sequence, mature HER-2 membrane distal extracellular and
XX intracellular domains and a C-terminal tag
XX
XX SQ Sequence 1665 BP; 330 A; 568 C; 465 G; 302 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,87e-140 Length: 1665
Score: 3020.00 Matches: 555
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-821-883-1 (1-555) x RAD21564 (1-1665)

Qy 1 MetArgAlaProLeuLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPhe 20
Db 1 ATGAGAGCTGCACCCCTCCTCTGCGCAGGCGCAGCAAGCCCTTAGCCCTTGGCTTCTTGT 60

Qy 21 LeuLeuPhePheTLeuLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAla 40
Db 61 CTGCTTTTCTGGCTAGACCGAAGTGTAAGTCCAGGAGTGGCGCGCGCGCG 120

Qy 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
Db 121 TCGACCCAGTGTGCACCGCAGACAGATGAGCTGGGCTCCCTGCCAGTCCCGAGACC 180

Qy 61 HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlnAsnLeuGlu 80
Db 181 CACCTGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGGTGCAGGGAACCTGGAA 240

Qy 81 LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGlnValGln 100
Db 241 CTCACCTACCTGCCACCAATGCCAGCTGTGCTTCTCTGAGGATATCCAGGAGGTGCAG 300

Qy 101 GlyTyrValLeuLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
Db 301 GGCTACGTGCTCATCGCTCACACCAAGTGGGCGAGTCCCACTGCAGAGGCTGGGGATT 360

Qy 121 ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 140
Db 361 GTGCGAGGCCCCAGCTCTTTTGAGGACAACATATGCCCCCTGGCGTGTAGACAAATGGAGAC 420

```

```

Qy 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln 160
Db 421 CCCTGAGACATACACCCCTGTCTACAGGGGCTCCCCAGAGGCTCGGGAGCTGCAG 480

Qy 161 LeuArgSerLeuThrGluLeuLeuLysGlyValLeuLeuGlnArgAsnProGlnLeu 180
Db 481 CTTTGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGATCAGCGGAACCCAGCTC 540

Qy 181 CysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnGlnLeuAlaLeu 200
Db 541 TGTACAGGACACGATTTTGTGAAGGACATCTCCACAGAACCAACAGCTGGCTCTC 600

Qy 201 ThrLeuLeuAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220
Db 601 AACTGATAGACACCAACCCCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGTAAGGC 660

Qy 221 SerArgCysTrpGlyGluSerGluAspCysGlnSerLeuThrArgThrValCysAla 240
Db 661 TCCCGCTGCTGGGGAGAGATTCTGAGGATTGTGAGAGCCTGACGCGCACTGTCTGTGCC 720

Qy 241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCysAla 260
Db 721 GGTGGCTGTGCCCTGCAAGGGGCGCACTGCCCACTGACTGCTGCCATGAGCACTGTGCT 780

Qy 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
Db 781 GCCGGCTGCAGGGGCGCCAGCACTCTGACTGCTGGCTGCCCTCCACTTCAACCCAGT 840

Qy 281 GlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSer 300
Db 841 GGCATCTGTGAGCTGCACCTGCCAGGCCCTGGTCACTACCAACACAGACAGCTTGTAGTCC 900

Qy 301 MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr 320
Db 901 ATGCCCAATCCGAGGGCGGTATACATTCGGGCGCAGCTGTGTGACTGCTGTGCTCTAC 960

Qy 321 AsnTyrLeuSerThrAspValGlySerGlyValAlaGlyValMetValHisHisArgHisArg 340
Db 961 AACTACTCTTCTACGACGCTGGATCGGGGCTGGGGGCGATGTTCCACACAGACACCGC 1020

Qy 341 SerSerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGlu 360
Db 1021 AGCTCATCTACAGAGGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAG 1080

Qy 361 GluAlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGly 380
Db 1081 GAGGCCCCAGGCTTCCACTGGCACCCCTCGAAGGGGCTGGCTCCGATGTATTTGATGGT 1140

Qy 381 AspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerPro 400
Db 1141 GACTGGGATGGGGGCGCAGCAAGGGCTGCAAAAGCTCCCCACACATAGACCCAGCCCT 1200

Qy 401 LeuGlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrVal 420
Db 1201 CTACAGCGGTACAGTGAAGGCCACAGTACCCCTGCCCTCTGAGACTGATGGTACGTT 1260

Qy 421 AlaProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGln 440
Db 1261 GCCCCCTTGACCTGCAGCCCCCAGCTGAAATATGTGAACCAAGCAGATGTTTCGGCCCCAG 1320

Qy 441 ProProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlu 460
Db 1321 CCCCTCTCGCCCCAGAGGGCCCTCTGCTGCTGCCGACCTGCTGGTGCACCTCTGGAA 1380

Qy 461 ArgAlaLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly 480
Db 1381 AGGGCCAAAGACTCTCTCCCGAGGGAAGATGGGGTGGTCAAAGACGTTTTTGGCTTTGGG 1440

Qy 481 GlyAlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHis 500
Db 1441 GGTGGCTGGAGAACCCCGAGTACTTGACACCCCGAGGAGGAGCTGCCCTCTAGCCCCAC 1500

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:41:11 ; Search time 5945.99 Seconds
(without alignments)
4045.651 Million cell updates/sec

Title: US-09-821-883-1
Perfect score: 3020
Sequence: 1 MRAAPLLAARASLSGLFLP.....ENPEYLGLDVPAAAHHHHH 555

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO/US09821883/runat_09092004_105127_6631/app_query.fasta_1.3100
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0 -LOPEXT=0 -LOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09821883 @CGN 1 1 16795 @runat_09092004_105127_6631 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_ats:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: gb_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_ats:*
27: em_un:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3020	100.0	1665	6	AX268285 Sequence
2	3005.5	99.5	1692	6	AX268287 Sequence
3	2972	98.4	2070	6	AX268286 Sequence
4	2957.5	97.9	2091	6	AX268288 Sequence
5	2416	80.0	4530	6	I21124 Sequence 9
6	2416	80.0	4530	6	I59745 Sequence 9
7	2416	80.0	4530	6	AR202597 Sequence
8	2416	80.0	4530	6	AR283481 Sequence
9	2416	80.0	4530	6	AR344811 Sequence
10	2416	80.0	4530	6	AX282577 Sequence
11	2416	80.0	4530	6	AX587649 Sequence
12	2416	80.0	4530	6	AX644071 Sequence
13	2416	80.0	4530	6	AX771418 Sequence
14	2416	80.0	4530	6	BD005474 Cellular
15	2416	80.0	4530	9	M1730 Human tyros
16	2413.5	79.9	9274	6	AR409602 Sequence
17	2413.5	79.9	9274	6	AX060703 Sequence
18	2411	79.8	4473	6	AR080259 Sequence
19	2411	79.8	4473	6	AR167390 Sequence
20	2411	79.8	4473	6	AR392088 Sequence
21	2411	79.8	4473	9	X03363 Human c-erb
22	2410.5	79.8	3768	6	AR409603 Sequence
23	2410.5	79.8	3768	6	AX060704 Sequence
24	2410.5	79.8	3768	6	AX467229 Sequence
25	2410.5	79.8	3768	6	AX481438 Sequence
26	2410.5	79.8	3768	6	BD224136 Novel met
27	2405.5	79.7	3768	6	AR034479 Sequence
28	2405.5	79.7	3768	6	BD267514 HER-2/neu
29	2405.5	79.7	3768	6	AX201817 Sequence
30	2405.5	79.7	3768	6	AX380923 Sequence
31	2405.5	79.7	3768	6	AX384604 Sequence
32	2405.5	79.7	3768	6	AX465456 Sequence
33	2334.5	77.3	3678	6	AX505114 Sequence
34	2129	70.5	2763	6	AX380942 Sequence
35	2129	70.5	2781	6	AX380944 Sequence
36	2124.5	70.3	3780	4	AB008451 Canis fam
37	1977	65.5	4506	10	AK129487 Mus muscu
38	1963.5	65.0	4694	10	BC046811 Mus muscu
39	1963.5	65.0	4695	10	BC053078 Mus muscu
40	1961	64.9	3771	6	BD267516 HER-2/neu
41	1961	64.9	3771	6	AX189662 Sequence
42	1961	64.9	3771	6	AX380925 Sequence
43	1958	64.8	4062	10	HAMNEU
44	1957	64.8	4727	10	AY116182
45	1956.5	64.8	3955	6	BD267515 HER-2/neu

ALIGNMENTS

RESULT 1

AX268285	AX268285	1665 bp	DNA	linear	PAT 29-OCT-2001
LOCUS	Sequence 6 from Patent WO0174855.				
DEFINITION	AX268285				
ACCESSION	AX268285.1	GI:16541537			
VERSION					
KEYWORDS	synthetic construct				
SOURCE	artificial sequences.				
ORGANISM					
REFERENCE	1				
AUTHORS	Laus, R., Vidovic, D. and Graddis, T.				
TITLE	Compositions and methods for dendritic cell-based immunotherapy				
JOURNAL	Patent: WO 0174855-A 6 11-OCT-2001;				
FEATURES	DENDREON CORPORATION (US)				
source	Location/Qualifiers				
	1..1665				
	/organism="synthetic construct"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:32630"				
	/note="HER500 construct"				
ORIGIN					
Alignment Scores:					
Pred. No.:	2.44e-110	Length:	1665		
Score:	3020.00	Matches:	555		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		
US-09-821-883-1 (1-555) x AX268285 (1-1665)					
Qy	1	MetArgAlaLapProLeuLeuAlaArgAlaLaserLeuSerLeuGlyPheLeuPhe	20		
Db	1	ATGAGAGCTGCACCCCTCCTCTGGCCAGGCGACAGCCCTTAGCCTTGGCTTCTTGGTT	60		
Qy	21	LeuLeuPhePheTtpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyValaAla	40		
Db	61	CTGCTTTTCTGGCTAGACCGAAGTGTACTACCAAGGAGTTGGCGGGGGCGCG	120		
Qy	41	SerThrGlnValCysThrGlyThrAspMetIysLeuArgLeuProAlaSerProGluThr	60		
Db	121	TCGACCAAGTGTGCACCGGCACAGACATGAAGCTGGCGCTCCCTGCGAGACC	180		
Qy	61	HisLeuAspMetLeuArgHisLeuTyrlGlnGlyCysGlnValValGlnGlnAsnLeu	80		
Db	181	CACCTGGACATGTCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGCAGGAAACCTGGAA	240		
Qy	81	LeuThrTyrlLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGln	100		
Db	241	CTACCTACCTGCCCCACCAATGCCAGCTGTCTCTCTGAGGATATCCAGGAGTGCAG	300		
Qy	101	GlyTyrlValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle	120		
Db	301	GGCTACGTGCTCATCGCTCAACAAGTGAAGGAGGCTCCACCTGCAGAGGCTCGGATT	360		
Qy	121	ValArgGlyThrGlnLeuPheGluAspAsnTyrlAlaLeuAlaValLeuAspAsnGlyAsp	140		
Db	361	GTGGAGGACCCAGCTCTTTGAGGACAACTATATCCCTGGCGGTGTAGCAATGGAGAC	420		
Qy	141	ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln	160		
Db	421	CCGCTGAACATACACCCCTGTACAGGGGCTCCCGAGGAGGCTCGGGAGCTGCAG	480		
Qy	161	LeuArgSerLeuThrGluLeuLeuLeuGlyGlyValLeuIleGlnArgAsnProGlnLeu	180		
Db	481	CTTCAAGACCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAAACCCAGCTC	540		
Qy	181	CysTyrlGlnAspThrIleLeuTtpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu	200		
Db	541	TGCTACCAAGGACATTTGTGAAGGACATCTTCCACAGAAACACACAGCTGGCTCTC	600		
Qy	201	ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly	220		
Db	601	ACACTGATAGACACCAACCCCTCTCGGGCCTGCACCCCTGTCTCCGATGTGTAAAGGC	660		
Qy	221	SerArgCysTtpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAla	240		
Db	661	TCCCGCTGTGGGGAGAGAGTTCTGAGGATTTGAGAGCTGTGAGCGCAGCTGTCTGTGCC	720		
Qy	241	GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCysAla	260		
Db	721	GGTGGCTGTCCCGCTGCAAGGGGCCACTGCCCCACTGCTGTGCTGAGCAGTGTGCT	780		
Qy	261	AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer	280		
Db	781	GCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGCTGCTCCCTCAACACAGT	840		
Qy	281	GlyIleCysGluLeuHisCysProAlaLeuValThrTyrlAsnThrAspThrPheGluSer	300		
Db	841	GGCATCTGTGAGCTGCACCTGCCAGCCCTGGTTCACCTACCAACACAGACACGTTTGA	900		
Qy	301	MetProAsnProGluGlyArgTyrlThrPheGlyAlaSerCysValThrAlaCysProTyrl	320		
Db	901	ATGCCCAATCCAGGGCCGTTATACATTCGGGCCAGCTGTGTGACTGCTGTCCCTAC	960		
Qy	321	AsnTyrlLeuSerThrAspValGlySerGlyAlaGlyGlyMetValHisHisArgHisArg	340		
Db	961	AACTACCTTTCTACGGACGTGGGATCGGGCGCTGGGGGCAATGGTCCACACAGCACCGC	1020		
Qy	341	SerSerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGlu	360		
Db	1021	AGCTCATCTACAGGAGTGGCGGTGACCTGACCTAGGGCTGGAGCCCTCTGAAGAG	1080		
Qy	361	GluAlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGly	380		
Db	1081	GAGGCCCCAGGTCTCCACTGGCACCTCCGAAGGGGCTGGCTCCGATGTATTGTATGGT	1140		
Qy	381	AspLeuGlyMetGlyAlaLysGlyLeuGlnSerLeuProThrHisAspProSerPro	400		
Db	1141	GACCTGGGAATGGGGCAGCAAGGGCTGCAAAAGCTCCCCACACATGACCCCGCCCT	1200		
Qy	401	LeuGlnArgTyrlSerGluAspProThrValProLeuProSerGluThrAspGlyTyrlVal	420		
Db	1201	CTACAGCGTACAGTGAAGACCCACAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1260		
Qy	421	AlaProLeuThrCysSerProGlnProGluTyrlValAsnGlnProAspValArgProGln	440		
Db	1261	GCCCCCTGACCTGACGCCCCCGAGCTGAATATGTGAACACAGCAGATGTTCGGCCCCAG	1320		
Qy	441	ProProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlu	460		
Db	1321	CCCCCTTGGCCCGAGAGGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1380		
Qy	461	ArgAlaLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly	480		
Db	1381	AGGGCCAAAGACTCTCTCCCGAGGGAAGATGGGGTCTGCTCAAGAGCTTTTGTGCTTTGGG	1440		
Qy	481	GlyAlaValGluAsnProGluTyrlLeuThrProGlnGlyGlyAlaAlaProGlnProHis	500		
Db	1441	GGTGGCTGTGAAGAACCCCGAGTACTTGAACACCCCGAGGAGAGCTGCCCCCTCAGCCCAC	1500		
Qy	501	ProProProAlaPheSerProAlaPheAspAsnLeuTyrlTyrlAspGlnAspProPro	520		
Db	1501	CCTCTCTCTGCTTACGCCCGAGCCCTTGCACAACTCTTATTACTGGGACAGGACCCACCA	1560		
Qy	521	GluArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyrl	540		
Db	1561	GAGCGGGGGCTCCACCCAGCCTTCAAGGGGACCTACGGCAGAGAAGACCCAGAGTAC	1620		
Qy	541	LeuGlyLeuAspValProAlaAlaHisHisHisHisHisHisHisHisHisHisHis	555		
Db	1621	CTGGGCTGTGACGTGCCCGCGCCGACATCACCATCACCATCACCATCACCATCACCAT	1665		
RESULT 2					
AX268287					

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:28:07 ; Search time 55.1456 Seconds
(without alignments)
3175.456 Million cell updates/sec

Title: US-09-821-883-1

Perfect score: 3020

Sequence: 1 MRAAPLLAARASLSGLF.....ENPEYGLDVPAAHFHHH 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organella.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2124.5	70.3	1259	6	O18735 canis fam
2	1951	64.6	1259	11	Q8K3F9
3	1609.5	53.3	419	4	Q9UK79
4	1376.5	45.6	711	11	Q80Y89
5	1005.5	33.3	881	11	Q8C0E7
6	969	32.1	367	11	Q8R2X1
7	906	30.0	165	4	Q14256
8	900.5	29.8	412	4	Q8WYV0
9	861.5	28.5	431	13	Q7SY19
10	754.5	25.0	527	13	Q90836
11	723	23.9	149	6	Q9BG66
12	717.5	23.8	643	11	Q9ERV6
13	717.5	23.8	655	11	Q9WVF5
14	717.5	23.8	1210	11	Q9EP98
15	714.5	23.7	1209	11	Q9QX70
16	708.5	23.5	1209	6	Q8MIL8

ALIGNMENTS

RESULT 1

ID	O18735	PRELIMINARY;	PRT; 1259 AA.
AC	O18735;		
DT	01-JAN-1998	(TrEMBLrel. 05, Created)	
DT	01-JAN-1998	(TrEMBLrel. 05, Last sequence update)	
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)	
DE	Erbb-2		
OS	Canis familiaris (Dog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.		
OX	NCBI_TaxID=9615;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Yokota H.;		
RT	"CDNA cloning of erbb-2 from canine mammary gland.";		
RL	Submitted (OCT-1997) to the EMBL/GenBank/DBSJ databases.		
DR	EMBL; AB008451; BAA23127.1; -		
DR	HSSP; P11362; 1FGK		
DR	GO; GO:0016020; C:membrane; IEA.		
DR	GO; GO:0005534; F:ATP binding; IEA.		
DR	GO; GO:0005509; F:calcium ion binding; IEA.		
DR	GO; GO:0005006; F:epidermal growth factor; IEA.		
DR	GO; GO:0016740; F:transferase activity; IEA.		
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.		
DR	GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IEA.		
DR	InterPro; IPR002048; EF-hand.		
DR	InterPro; IPR006211; Furin-like.		
DR	InterPro; IPR006212; Furin repeat.		
DR	InterPro; IPR009030; Grow_Fac_recep.		
DR	InterPro; IPR000719; Prot_kinase.		
DR	InterPro; IPR001245; Tyr_kinase.		
DR	InterPro; IPR008266; Tyr_kinase_AS.		
DR	InterPro; IPR004019; VLP motif.		
DR	Pfam; PF00757; Furin-like; 1.		
DR	Pfam; PF00065; pkinase; 1.		
DR	Pfam; PF01030; Recep_L_domain; 2.		
DR	Pfam; PF02757; VLP; 2.		

Q9ESE0 rattus norv
Q78ZF7 brachydanio
Q9BUD7 homo sapien
Q9YH40 xiphophorus
Q8AW81 brachydanio
Q9PEH2 gallus gall
P79754 fugu rubrip
Q8MIW0 drosophila
Q9BIH9 anopheles g
Q86NZ2 drosophila
Q9W6F6 gallus gall
Q8BG64 oryctolagus
Q9BGE5 oryctolagus
Q23821 caenorhabdi
Q26569 schistosoma
Q26566 schistosoma
Q26567 schistosoma
Q26568 schistosoma
Q9YLX8 ephydatia f
Q8EMD7 echinococcu
Q9USA8 bombyx mori
Q9NJVS biomphalari
Q8TOW6 echinococcu
Q88459 mus musculu
Q8UW83 paralichthy
Q93457 scophthalmu
Q9VD94 drosophila
Q73798 xenopus lae
Q8UW84 paralichthy

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:20:54 ; Search time 10.1875 Seconds
(without alignments)
2836.691 Million cell updates/sec

Title: US-09-821-883-1

Perfect score: 3020

Sequence: 1 MRAAPLLARASLSGLFLP.....ENPEYGLDVPAAHHHHH 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2405.5	79.7	1255	1 ERB2 HUMAN	P04626 homo sapien
2	1947.5	64.5	1257	1 ERB2 RAT	P06494 rattus norv
3	1935	64.1	1254	1 ERB2 MESAU	Q60553 mesocricetu
4	755.5	25.0	703	1 EGFR CHICK	P13387 gallus gall
5	721	23.9	1210	1 EGFR HUMAN	P00533 homo sapien
6	721	23.9	1308	1 ERB4 HUMAN	Q15303 homo sapien
7	717.5	23.8	1210	1 EGFR MOUSE	Q01279 mus musculu
8	717	23.7	1308	1 ERB4 RAT	Q62956 rattus norv
9	697	23.1	1342	1 ERB3 HUMAN	P21860 homo sapien
10	668	22.1	1339	1 ERB3 RAT	Q62799 rattus norv
11	655.5	21.7	1167	1 XMRK XIPMA	P13388 xiphophorus
12	536	17.7	1426	1 EGFR DROME	P04412 drosophila
13	404.5	13.4	1367	1 L723 CAEEL	P24348 caenorhabdi
14	317.5	10.5	1363	1 ILPR BRAUA	O02466 bruchioisto
15	293	9.7	245	1 ERB2 MOUSE	P70424 mus musculu
16	273.5	9.1	1477	1 HTK7 HYDAT	Q25197 hydra atten
17	270.5	9.0	2146	1 INSR DROME	P09208 drosophila
18	256.5	8.5	1382	1 INSR HUMAN	P06213 homo sapien
19	232.5	8.4	1372	1 INSR MOUSE	P15208 mus musculu
20	252	8.3	1607	1 MPR LYMSY	Q25410 lymphnaea sta
21	251.5	8.3	1300	1 IRR MOUSE	Q9wt14 mus musculu
22	249	8.2	1383	1 INSR RAT	P15127 rattus norv
23	248.5	8.2	1297	1 IRR HUMAN	P14616 homo sapien
24	246.5	8.2	1300	1 IRR CAVPO	P14617 cavia porce
25	246	8.1	1367	1 IGR HUMAN	P08069 homo sapien
26	235.5	7.8	1370	1 IGR RAT	P24062 rattus norv
27	233.5	7.7	1373	1 IGR MOUSE	Q60751 mus musculu
28	228.5	7.6	581	1 IRR RAT	Q64716 rattus norv
29	207	6.9	1390	1 INSR AEDAE	Q93105 aedes aegyp
30	168	5.6	386	1 PPAP HUMAN	P15309 homo sapien
31	155	5.1	1321	1 IRS2 MOUSE	P81122 mus musculu
32	151.5	5.0	707	1 SPQ HUMAN	P23246 homo sapien
33	151.5	5.0	5262	1 MLL2 HUMAN	O14686 homo sapien

RESULT 1
ERB2_HUMAN
ID ERB2_HUMAN STANDARD; PRT; 1255 AA.
AC P04626;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell surface receptor HER2) (MLN 19).
GN ERBB2 OR HER2 OR NGL OR NEU.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96118663; PubMed=3003577;
RA Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N., Saito T., Toyoshima K.;
RT "Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth factor receptor.";
RL Nature 319:230-234(1986).
[2]
RP SEQUENCE FROM N.A., AND VARIANT ALA-1170.
RX MEDLINE=86070181; PubMed=2999974;
RA Coussens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A., McGrath J., Seeburg P.H., Libermann T.A., Schlessinger J., Francke U., Levinson A., Ullrich A.;
RT "Tyrosine kinase receptor with extensive homology to EGF receptor shares chromosomal location with neu oncogene.";
RL Science 230:1132-1139(1985).
[3]
RP SEQUENCE FROM N.A., AND VARIANTS CYS-452; VAL-655 AND ALA-1170.
RX Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M.-W., Miyamoto K.E., Nguyen C.P., Nguyen D.A., Peol C.L., Robertson P.D., Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RT Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE OF 737-1031 FROM N.A.
RX MEDLINE=86016729; PubMed=295967;
RA Semba K., Kamata N., Toyoshima K., Yamamoto T.;
RT "A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epidermal growth factor-receptor gene and is amplified in a human salivary gland adenocarcinoma.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).
[5]
RP VARIANTS VAL-654 AND VAL-655.
RX MEDLINE=93194196; PubMed=8095488;
RA Ehsani A., Low J., Wallace R.B., Wu A.M.;
RT "Characterization of a new allele of the human ERBB2 gene by allele-specific competition hybridization.";
RL Genomics 15:426-429(1993).
CC -1- FUNCTION: Essential component of a neurotrophin-receptor complex, although neurotrophins do not interact with it alone. GP30 is a potential ligand for this receptor. Not activated by EGF, TGF-

ALIGNMENTS

34 141.5 4.7 830 1 SREC_HUMAN Q14162 homo sapien
35 141.5 4.7 2442 1 CBP_HUMAN Q92793 homo sapien
36 140 4.6 863 1 M113_HUMAN Q9n3f8 homo sapien
37 140 4.6 1696 1 PKC5_BRACL Q9n155 branchiosto
38 139 4.6 634 1 ERBB_ALV P00534 avian leuko
39 139 4.6 870 1 SRC2_HUMAN Q959p6 homo sapien
40 138.5 4.6 553 1 ODO2_MYCTU Q10381 mycobacteri
41 137.5 4.6 620 1 EXTN_TOBAC P13983 nicotiana t
42 136 4.5 331 1 PRP1_HUMAN P04280 homo sapien
43 133.5 4.4 886 1 SM6B_MOUSE O54951 mus musculu
44 131.5 4.4 1185 1 DRPL_HUMAN P54259 homo sapien
45 130.5 4.3 833 1 SRC2_MOUSE P59222 mus musculu

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:29:10 ; Search time 17.7175 Seconds
(without alignments)
3013.200 Million cell updates/sec

Title: US-09-821-883-1
Perfect score: 3020
Sequence: 1.MRAAPLLARASLSLGLF.....ENPEYGLDVPAAAHHHHHH 555
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2405.5	79.7	1255	1 A24571	protein-tyrosine k
2	1950.5	64.6	1260	1 TVRTNU	protein-tyrosine k
3	1935	64.1	1254	2 I48161	p-185 precursor -
4	755.5	25.0	1223	1 TVCHLV	epidermal growth f
5	754.5	25.0	527	2 A42032	epidermal growth f
6	721	23.9	1210	1 GQHUE	epidermal growth f
7	721	23.9	1308	2 A47253	epidermal growth f
8	717.5	23.8	1210	2 A53183	epidermal growth f
9	714.5	23.7	644	2 A36325	epidermal growth f
10	697	23.1	1342	2 A36223	kinase-related tra
11	662	21.9	1339	2 J4387	epidermal growth f
12	655.5	21.7	1166	1 S06142	protein-tyrosine k
13	536	17.7	843	2 A27131	epidermal growth f
14	404.5	13.4	1323	2 E88257	protein let-23 (im
15	404.5	13.4	1374	2 S70712	protein-tyrosine k
16	390.5	12.9	1369	2 S70713	protein-tyrosine k
17	375.5	12.4	1330	1 GQFFE	epidermal growth f
18	341	11.3	366	2 D45558	epidermal growth f
19	341	11.3	1717	1 A45558	epidermal growth f
20	331	11.0	333	2 B45558	epidermal growth f
21	331	11.0	342	2 C45558	epidermal growth f
22	317.5	10.5	1363	2 T43220	insulin-like growth
23	273.5	9.1	1477	2 T18534	protein-tyrosine k
24	270.5	9.0	2101	2 S57245	insulin receptor (
25	270.5	9.0	2148	1 A56081	insulin receptor -
26	256.5	8.5	1382	1 INHUR	insulin receptor p
27	252.5	8.4	1372	2 A34157	insulin receptor p
28	252	8.3	1607	2 T43212	insulin-like growth
29	249	8.2	1383	2 A36080	insulin receptor p

30 246.5 8.2 1300 2 A36502 insulin receptor-r
31 246 8.1 1367 1 IGHURI insulin-like growth
32 238 7.9 540 2 B47417 insulin receptor-r
33 235.5 7.8 183 2 JH0803 tyrosine kinase re
34 235.5 7.8 1371 2 A33837 insulin-like growth
35 231.5 7.7 1268 2 B36502 insulin receptor-r
36 231.5 7.4 329 2 A48805 insulin-like growth
37 207 6.9 1390 2 T30346 insulin receptor -
38 188 6.2 1846 2 T42047 insulin receptor h
39 168 5.6 386 1 JH0610 acid phosphatase (

ALIGNMENTS

RESULT 1

A24571
protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human
N:Alternate names: C-erb-B-2 protein precursor; kinase-related transforming protein erbB
C:Species: Homo sapiens (man)
C:Date: 25-Oct-1987 #sequence revision 06-Dec-1996 #text_change 11-Jun-1999
C:Accession: A24571; A25491; A44188; B44188; I59509; I57622
R:Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T.;
Nature 319, 230-234, 1986
A:Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth f
A:Reference number: A24571; MUID:86118663; PMID:3003577
A:Accession: A24571
A:Molecule type: mRNA
A:Residues: 1-1255 <YAM>
A:Cross-references: GB:X03363; NID:g31197; PIDN:CAA27060.1; PID:g31198
R:Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985
A:Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epiderm
A:Reference number: A25491; MUID:86016729; PMID:2995967
A:Accession: A25491
A:Molecule type: DNA
A:Residues: 737-1031 <SEM>
A:Cross-references: GB:M11767; NID:g182163; PIDN:AAA35808.1; PID:g553282
R:Coussens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg, P.
Science 230, 1132-1139, 1985
A:Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chromos
A:Reference number: A44188; MUID:86070181; PMID:2999974
A:Accession: A44188
A:Molecule type: DNA
A:Residues: 740-910 <COU1>
A:Cross-references: GB:M12036; NID:g183988; PIDN:AAA35978.1; PID:g183989
A:Accession: B44188
A:Molecule type: mRNA
A:Residues: 1-517, 'RALL', 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>
A:Cross-references: GB:M11730; NID:g183986
R:King, C.R.; Kraus, M.H.; Aaronson, S.A.
Science 229, 974-976, 1985
A:Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.
A:Reference number: I59509; MUID:85272597; PMID:2992089
A:Accession: I59509
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 832-909 <REX>
A:Cross-references: GB:I29395; NID:g459807; PIDN:AAA35809.1; PID:g459808
R:Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.
Mol. Cell. Biol. 7, 2597-2601, 1987
A:Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptional
A:Reference number: I57622; MUID:87286898; PMID:3039351
A:Accession: I57622
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-191 <TAL>

A;Cross-references: GB:M16792; NID:g183983; PIDN:AAA58637.1; PID:g553332
C;Comment: Amplification and overexpression of this erbB-related gene occurs in about 30
C;Genetics:
A;Gene: GDB:ERBB2; NGL; NEU; HER-2
A;Cross-references: GDB:120613; OMIM:164870
A;Map position: 17q21.1-17q21.1
A;Introns: 25/1; 75/3; 147/1; 883/3
A;Note: the list of introns is incomplete
C;Function:
A;Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
A;Superfamily: epidermal growth factor receptor; protein kinase homology
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
inase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-1255/Product: protein-tyrosine kinase erbB2 #status predicted <MAT>
F;22-653/Domain: extracellular #status predicted <EXT>
F;70-304/Domain: EGF receptor extracellular domain repeat <EB1>
F;395-605/Domain: EGF receptor extracellular domain repeat <EB2>
F;654-675/Domain: transmembrane #status predicted <TMW>
F;676-1255/Domain: intracellular #status predicted <INT>
F;718-983/Domain: protein kinase homology <KIN>
F;726-734/Region: protein kinase ATP-binding motif
F;68,124,187,259,530,571,629/Binding site: carbohydrate (Asn) (covalent) #status predict
F;686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F;753/Active site: Lys #status predicted
F;1139,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 79.7%; Score 2405.5; DB 1; Length 1255;
Best Local Similarity 41.1%; Pred. No. 3.7e-133;
Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

QY	35	LARGAATQVCTGDMKRLPASPTHLDMLRHLYQGVQVQGNLELYLPTNASLSFLQ	94
DB	16	LPPGAATQVCTGDMKRLPASPTHLDMLRHLYQGVQVQGNLELYLPTNASLSFLQ	75
QY	95	DIQVQGVVLIHQNVRQVPLQRLIRVGTQLFEDNVALAVLDNGDPLNNTPTVTGASPG	154
DB	76	DIQVQGVVLIHQNVRQVPLQRLIRVGTQLFEDNVALAVLDNGDPLNNTPTVTGASPG	135
QY	155	GLRELQLRSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNQALALIDTNRSRACHPC	214
DB	136	GLRELQLRSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNQALALIDTNRSRACHPC	195
QY	215	SPMCKGRRCWGESSEDCQSLTRITVCAGGCARCKGFLPTDCHEQCAAGCTGPKHSDCLAC	274
DB	196	SPMCKGRRCWGESSEDCQSLTRITVCAGGCARCKGFLPTDCHEQCAAGCTGPKHSDCLAC	255
QY	275	LHFHNSGICELHCPALVTYNTDTFESMPNPEGRVTFGASCVTACFYNYLSTDVGS----	329
DB	256	LHFHNSGICELHCPALVTYNTDTFESMPNPEGRVTFGASCVTACFYNYLSTDVGSCTLVC	315
QY	330	-----	329
DB	316	PLHNQVETAEADGTQRCCKSPCARVCYGLGMEHLREVRVTSANIQEFAGCKKIFGSLA	375
QY	330	-----	329
DB	376	FLPESFDGPASNTAPLQEQVETLEITGYLYISAWPDSLPLDSVFQNLQVIRGRI	435
QY	330	-----	329
DB	436	LHNGAYSLTLOGLIGLSWGLRSLRSLBELSGSLALIHNTLHCFVHTVPWDLFPNPHQALLH	495
QY	330	-----	329
DB	496	TANRPEBCEVGEGLACHQLCARGHCWGPQTQVCNCSQFLRGQECVECRVLQGLPREYV	555
QY	330	-----	329
DB	556	NARHCLPCHPECQPNQSVTCFGEPAQCVAHYKDPFPCVARCPGSKVGPDLSYMPTWK	615
QY	330	-----	329

DB	616	FPDEGACPCPINCTHSCVDLDDKGCAPAEQASPLTISIISAVVGLLVVVLGVVFGILI	675
QY	330	-----	329
DB	676	KRFQOKIRKYTMRLLIQETELVEPLTFPSGAMPNQAOIRILKETELRKVKVLGSGAFGVY	735
QY	330	-----	329
DB	736	KGIWIPDGENVKIPVALKVLRENTSPKANKEILDEAVMAGVGSPPYVRLIGLICLTSTVQ	795
QY	330	-----	329
DB	796	LVTQLMPYGLLDHVRNRRGLSGQDLLNWCMIQAKGMSYLEVDVRLVHRDLAARNVLVKS	855
QY	330	-----	329
DB	856	PNHVKITDFGLARLLDIDETEHADGGKVPKWMALLEIILRRRFTHQSDVMSYGVTTWEL	915
QY	330	-----	329
DB	916	MTFGAKPYDGIIPAREIPDLLLEKGERLPQPPICTIDVTVMVKWMIDSECRPRPRELVE	975
QY	330	-----	329
DB	976	FSRMARDPQRFVVIQNEEDLGPASPLDSTFYRSLLDDDDMGDLVDABEYLVPOQGFCDP	1035
QY	330	-----	387
DB	1036	APGAGGVHHRSSSTRSGGGDLTLGLEPSEAEPRSLAPSEGAGSDVFDGDLNGMAA	1095
QY	388	KGLQSLPTHTDPSLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPQVVRPQPPSPREG	447
DB	1096	KGLQSLPTHTDPSLQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPQVVRPQPPSPREG	1155
QY	448	PLPAARPAAGATLBRKATLSFGKNGVVDVAFGAGAVENPEYLTPOGGAAAPQHPPPAFSP	507
DB	1156	PLPAARPAAGATLBRKATLSFGKNGVVDVAFGAGAVENPEYLTPOGGAAAPQHPPPAFSP	1215
QY	508	AFNLVYWDQDPPERGAPPSTFKGTPTAENPEYLGLDVP	546
DB	1216	AFNLVYWDQDPPERGAPPSTFKGTPTAENPEYLGLDVP	1254

RESULT 2

TVRTNU
protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 11-Jun-1999
C;Accession: A24562; A61204
R;Bargmann, C.I.; Hung, M.C.; Weinberg, R.A.
Nature 319, 226-230, 1986
A;Title: The neu oncogene encodes an epidermal growth factor receptor-related protein.
A;Reference number: A24562; UID:86118662; PMID:3945311
A;Accession: A24562
A;Molecule type: mRNA
A;Residues: 1-1260 <BAR>
A;Cross-references: EMBL:X03362; NID:g56745; PIDN:CAA27059.1; PID:g56746
R;Masui, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohen, R.; Carcinogenesis 12, 1975-1978, 1991
A;Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals no m-2-thiazolyl]formamide or N-methyl-N-nitrosourea.
A;Reference number: A61204; UID:92035293; PMID:1682063
A;Accession: A61204
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 637-663, 'V', 665-702 <MAS>
A;Note: authors translated the codon GCA for residue 25 as Val
C;Genetics:
A;Gene: neu
C;Superfamily: epidermal growth factor receptor; protein kinase homology
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:41:37 ; Search time 58.6891 Seconds
(without alignments)
3032.635 Million cell updates/sec

Title: US-09-821-883-1

Perfect score: 3020

Sequence: 1 MRAAPLLARASLSLGFLF.....ENPEYGLDVPAAHHHHH 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3020	100.0	555	US-09-821-883-1	Sequence 1, Appli
2	3005.5	99.5	564	US-09-821-883-3	Sequence 3, Appli
3	2972	98.4	690	US-09-821-883-2	Sequence 2, Appli
4	2957.5	97.9	697	US-09-821-883-4	Sequence 4, Appli
5	2573.5	85.2	919	US-09-854-356-6	Sequence 6, Appli
6	2410.5	79.8	1255	US-09-811-123-9	Sequence 9, Appli
7	2410.5	79.8	1255	US-09-811-115-3	Sequence 3, Appli
8	2410.5	79.8	1255	US-09-769-508-2	Sequence 2, Appli
9	2410.5	79.8	1255	US-09-984-092-4	Sequence 4, Appli
10	2410.5	79.8	1255	US-10-177-293-126	Sequence 126, App
11	2410.5	79.8	1255	US-10-207-498-6	Sequence 6, Appli
12	2410.5	79.8	1255	US-10-338-730-2	Sequence 2, Appli
13	2410.5	79.8	1255	US-10-322-892-4	Sequence 4, Appli
14	2410.5	79.8	1255	US-10-272-437A-28	Sequence 28, Appli
15	2410.5	79.8	1255	US-10-117-937-594	Sequence 594, App

ALIGNMENTS

RESULT 1

US-09-821-883-1
; Sequence 1, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; TITLE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500 construct
US-09-821-883-1

Query Match 100.0%; Score 3020; DB 9; Length 555;
Best Local Similarity 100.0%; Pred. No. 4.5e-203;
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAAPLLARASLSLGFLFLLFFWLDKRSVLAKELARGAASQVCTGTGDKMLRASPPT 60
DB 1 MRAAPLLARASLSLGFLFLLFFWLDKRSVLAKELARGAASQVCTGTGDKMLRASPPT 60
QY 61 HLDMLRHLYQGCQVQVGNLELYLPTNASLSFLQDIEQVGVVLAHNOVROVPLQRLRI 120
DB 61 HLDMLRHLYQGCQVQVGNLELYLPTNASLSFLQDIEQVGVVLAHNOVROVPLQRLRI 120
QY 121 VRGTQIFEDNYALAVLDNGDPLNNTTPTVTGASFGGLRELQRLSLTEILKGGVLIQRNPOL 180

16 2410.5 79.8 1255 15 US-10-435-696-36 Sequence 36, Appli
17 2410.5 79.8 1255 16 US-10-441-779C-4 Sequence 4, Appli
18 2410.5 79.8 1255 16 US-10-734-564-126 Sequence 126, App
19 2405.5 79.7 1255 9 US-09-854-356-1 Sequence 1, Appli
20 2405.5 79.7 1255 10 US-09-930-125-2 Sequence 2, Appli
21 2405.5 79.7 1255 10 US-09-441-411-6 Sequence 6, Appli
22 2405.5 79.7 1255 12 US-10-463-162-3 Sequence 3, Appli
23 2405.5 79.7 1255 12 US-10-253-286-553 Sequence 553, App
24 2405.5 79.7 1255 12 US-09-765-973-2 Sequence 2, Appli
25 2405.5 79.7 1255 12 US-10-418-027-3 Sequence 3, Appli
26 2405.5 79.7 1255 14 US-10-207-655-45 Sequence 45, Appli
27 2405.5 79.7 1255 14 US-10-313-644-2 Sequence 2, Appli
28 2405.5 79.7 1255 15 US-10-394-322A-17 Sequence 17, Appli
29 2405.5 79.7 1255 15 US-10-245-871-553 Sequence 553, App
30 2405.5 79.7 1255 15 US-10-149-138-4641 Sequence 4641, App
31 2405.5 79.7 1255 16 US-10-647-005-68 Sequence 68, Appli
32 2405.5 79.7 1255 16 US-10-149-138-4641 Sequence 4641, App
33 2398.5 79.4 1253 14 US-10-146-473-72 Sequence 72, Appli
34 1961 64.9 1256 9 US-09-854-356-14 Sequence 14, Appli
35 1950.5 64.6 1260 9 US-09-870-759-118 Sequence 118, App
36 1950.5 64.6 1260 10 US-09-751-708A-118 Sequence 118, App
37 1948 64.5 1256 9 US-09-854-356-2 Sequence 2, Appli
38 1795 59.4 479 9 US-09-821-883-5 Sequence 5, Appli
39 1612 53.4 645 9 US-09-921-161-1 Sequence 1, Appli
40 1612 53.4 645 14 US-10-268-501-13 Sequence 13, Appli
41 1612 53.4 645 15 US-10-608-626-13 Sequence 13, Appli
42 1612 53.4 653 9 US-09-854-356-3 Sequence 3, Appli
43 1612 53.4 685 15 US-10-412-804A-4 Sequence 4, Appli
44 1612 53.4 690 15 US-10-412-804A-11 Sequence 11, Appli
45 1612 53.4 712 9 US-09-854-356-7 Sequence 7, Appli

Db 121 VGGTQLFEDNYALVLDNGDPLNNTTPTVGTASPGGLRELQLSLTEILKGGVLIQRNPOL 180
QY 181 CYQDTILWKDIFPHKNNQALTLIDTNRSRACHPCSPMCKGSRGWESSEDCCSLTRTVCA 240
Db 181 CYQDTILWKDIFPHKNNQALTLIDTNRSRACHPCSPMCKGSRGWESSEDCCSLTRTVCA 240
QY 241 GGCARCKGPLPTDCCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPES 300
Db 241 GGCARCKGPLPTDCCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPES 300
QY 301 MNPENGRYTFGASCVTACPYNYLSTDVSGAGGVMVHRRSSSTRSGGDLTLGLEPSEE 360
Db 301 MNPENGRYTFGASCVTACPYNYLSTDVSGAGGVMVHRRSSSTRSGGDLTLGLEPSEE 360
QY 361 EAPRSLAPSEAGSDVFDGLGMAAGLQSLPHTDPSPLORYSEDPVPLPSETDGVV 420
Db 361 EAPRSLAPSEAGSDVFDGLGMAAGLQSLPHTDPSPLORYSEDPVPLPSETDGVV 420
QY 421 APLTCSPOPEYVNDVVRPQPPSPREGPLPAARPAAGATLERAKTILSPGKGVVVDVFAFG 480
Db 421 APLTCSPOPEYVNDVVRPQPPSPREGPLPAARPAAGATLERAKTILSPGKGVVVDVFAFG 480
QY 481 GAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLVYWDQDPPPERGAPPSTFKGTPTAENPEY 540
Db 481 GAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLVYWDQDPPPERGAPPSTFKGTPTAENPEY 540
QY 541 LGLDVPAHAHHHHH 555
Db 541 LGLDVPAHAHHHHH 555

RESULT 2

US-09-821-883-3

; Sequence 3, Application US/09821883

; Patent No. US20020061310A1

; GENERAL INFORMATION:

; APPLICANT: Laus, Reiner

; APPLICANT: Vidovic, Damir

; APPLICANT: Graddis, Thomas

; TITLE OF INVENTION: Compositions and Methods for Dendritic

; TITLE OF INVENTION: Cell-Based Immunotherapy

; FILE REFERENCE: 7636-0022.30

; CURRENT APPLICATION NUMBER: US/09/821,883

; CURRENT FILING DATE: 2001-03-30

; PRIOR APPLICATION NUMBER: US 60/193,504

; PRIOR FILING DATE: 2000-03-30

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 564

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: HER500* construct

US-09-821-883-3

Query Match 99.5%; Score 3005.5; DB 9; Length 564;
Best Local Similarity 98.4%; Pred. No. 4.8e-202;
Matches 555; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

QY 1 MRAAPLLIARAASLSLGLFLLFFWLDNRSLAKELARGASTQVCTGTDMKRLPASPET 60
Db 1 MRAAPLLIARAASLSLGLFLLFFWLDNRSLAKELARGASTQVCTGTDMKRLPASPET 60
QY 61 HLDMLRHLVQGCQVQGNLELYLPTNASLSFLQDIQEVQGVVLAHNQVQVPLQRLRI 120
Db 61 HLDMLRHLVQGCQVQGNLELYLPTNASLSFLQDIQEVQGVVLAHNQVQVPLQRLRI 120
QY 121 VGGTQLFEDNYALVLDNGDPLNNTTPTVGTASPGGLRELQLSLTEILKGGVLIQRNPOL 180
Db 121 VGGTQLFEDNYALVLDNGDPLNNTTPTVGTASPGGLRELQLSLTEILKGGVLIQRNPOL 180

QY 181 CYQDTILWKDIFPHKNNQALTLIDTNRSRACHPCSPMCKGSRGWESSEDCCSLTRTVCA 240
Db 181 CYQDTILWKDIFPHKNNQALTLIDTNRSRACHPCSPMCKGSRGWESSEDCCSLTRTVCA 240
QY 241 GGCARCKGPLPTDCCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPES 300
Db 241 GGCARCKGPLPTDCCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPES 300
QY 301 MNPENGRYTFGASCVTACPYNYLSTDVGS-----GAGGMVHRRSSSTRSGGDL 351
Db 301 MNPENGRYTFGASCVTACPYNYLSTDVGSASIIINEFKLGAGGMVHRRSSSTRSGGDL 360
QY 352 TLGLEPSEEEAAPSRLAPSEAGSDVFDGLGMAAGLQSLPHTDPSPLORYSEDPVPLP 411
Db 352 TLGLEPSEEEAAPSRLAPSEAGSDVFDGLGMAAGLQSLPHTDPSPLORYSEDPVPLP 420
QY 412 LPSETDGVVAPLTCSPQPEYVNDVVRPQPPSPREGPLPAARPAAGATLERAKTILSPGKNG 471
Db 412 LPSETDGVVAPLTCSPQPEYVNDVVRPQPPSPREGPLPAARPAAGATLERAKTILSPGKNG 480
QY 472 VVKOVFAFGGAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLVYWDQDPPPERGAPPSTFKG 531
Db 472 VVKOVFAFGGAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLVYWDQDPPPERGAPPSTFKG 540
QY 532 TPTAENPEYLTGIDVPAHAHHHHH 555
Db 532 TPTAENPEYLTGIDVPAHAHHHHH 555
QY 541 TPTAENPEYLTGIDVPAHAHHHHH 564
Db 541 TPTAENPEYLTGIDVPAHAHHHHH 564

RESULT 3

US-09-821-883-2

; Sequence 2, Application US/09821883

; Patent No. US20020061310A1

; GENERAL INFORMATION:

; APPLICANT: Laus, Reiner

; APPLICANT: Vidovic, Damir

; APPLICANT: Graddis, Thomas

; TITLE OF INVENTION: Compositions and Methods for Dendritic

; TITLE OF INVENTION: Cell-Based Immunotherapy

; FILE REFERENCE: 7636-0022.30

; CURRENT APPLICATION NUMBER: US/09/821,883

; CURRENT FILING DATE: 2001-03-30

; PRIOR APPLICATION NUMBER: US 60/193,504

; PRIOR FILING DATE: 2000-03-30

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 690

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: HER500-hGM-CSF construct

US-09-821-883-2

Query Match 98.4%; Score 2972; DB 9; Length 690;
Best Local Similarity 100.0%; Pred. No. 1.4e-199;
Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAAPLLIARAASLSLGLFLLFFWLDNRSLAKELARGASTQVCTGTDMKRLPASPET 60
Db 1 MRAAPLLIARAASLSLGLFLLFFWLDNRSLAKELARGASTQVCTGTDMKRLPASPET 60
QY 61 HLDMLRHLVQGCQVQGNLELYLPTNASLSFLQDIQEVQGVVLAHNQVQVPLQRLRI 120
Db 61 HLDMLRHLVQGCQVQGNLELYLPTNASLSFLQDIQEVQGVVLAHNQVQVPLQRLRI 120
QY 121 VGGTQLFEDNYALVLDNGDPLNNTTPTVGTASPGGLRELQLSLTEILKGGVLIQRNPOL 180
Db 121 VGGTQLFEDNYALVLDNGDPLNNTTPTVGTASPGGLRELQLSLTEILKGGVLIQRNPOL 180
QY 181 CYQDTILWKDIFPHKNNQALTLIDTNRSRACHPCSPMCKGSRGWESSEDCCSLTRTVCA 240
Db 181 CYQDTILWKDIFPHKNNQALTLIDTNRSRACHPCSPMCKGSRGWESSEDCCSLTRTVCA 240

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:31:25 ; Search time 16.1672 Seconds
(without alignments)
1772.256 Million cell updates/sec

Title: US-09-821-883-1

Perfect score: 3020

Sequence: 1 MRAAPLLARAASLSGLFL.....ENPEYLGLDVPAAHHHHH 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCTUS-COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2410.5	79.8	1255	4	US-09-527-487-2
2	2410.5	79.8	1255	4	US-09-811-115-3
3	2405.5	79.7	1255	1	US-08-467-083-68
4	2405.5	79.7	1255	1	US-08-414-417B-68
5	2405.5	79.7	1255	2	US-08-486-348A-68
6	2405.5	79.7	1255	2	US-08-625-101-2
7	2405.5	79.7	1255	2	US-08-468-545B-68
8	2405.5	79.7	1255	2	US-08-356-786-2
9	2405.5	79.7	1255	3	US-08-466-680B-68
10	2405.5	79.7	1255	4	US-09-354-533-68
11	2385.5	79.0	1255	2	US-08-484-438-8
12	1612	53.4	782	2	US-09-146-283-4
13	1612	53.4	782	3	US-08-579-823A-4
14	1612	53.4	782	3	US-09-344-195-4
15	1607	53.2	419	4	US-09-630-155-2
16	1597	52.9	624	3	US-08-422-108-1
17	1597	52.9	624	4	US-08-422-734-1
18	1183	39.2	580	1	US-08-414-417B-69
19	1183	39.2	580	2	US-08-486-348A-69
20	1183	39.2	580	2	US-08-468-545B-69
21	1183	39.2	580	3	US-08-466-680B-69
22	1183	39.2	580	4	US-09-354-533-69
23	860	28.5	166	4	US-09-648-067A-1
24	721	23.9	644	1	US-08-336-708A-9
25	721	23.9	911	2	US-08-484-438-10
26	721	23.9	1058	2	US-08-484-438-4
27	721	23.9	1210	2	US-08-484-438-7

28	721	23.9	1210	2	US-08-475-035-4	Sequence 4, Appli
29	721	23.9	1308	2	US-08-484-438-2	Sequence 2, Appli
30	704.5	23.3	478	4	US-09-570-454-2	Sequence 2, Appli
31	704.5	23.3	478	4	US-09-867-521-2	Sequence 2, Appli
32	698.5	23.1	1343	6	5183884-4	Patent No. 5183884
33	697	23.1	1342	1	US-07-978-895-4	Sequence 4, Appli
34	697	23.1	1342	2	US-08-484-438-9	Sequence 9, Appli
35	697	23.1	1342	2	US-08-473-119-4	Sequence 4, Appli
36	697	23.1	1342	2	US-08-475-352-4	Sequence 4, Appli
37	697	23.1	1342	4	US-09-170-699-4	Sequence 4, Appli
38	493	16.3	97	1	US-08-421-356-3	Sequence 3, Appli
39	493	16.3	97	4	US-09-046-783-3	Sequence 3, Appli
40	257.5	8.5	1382	2	US-08-737-715-2	Sequence 2, Appli
41	257.5	8.5	1382	4	US-09-457-040B-7	Sequence 7, Appli
42	246	8.1	516	3	US-08-746-559A-4	Sequence 4, Appli
43	246	8.1	1367	2	US-08-249-687C-2	Sequence 2, Appli
44	246	8.1	1367	2	US-08-625-819-2	Sequence 2, Appli
45	246	8.1	1367	3	US-08-746-559A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-527-487-2

; Sequence 2, Application US/09527487

; Patent No. 6528060

; GENERAL INFORMATION:

; APPLICANT: Nicolette, Charles

; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES

; FILE REFERENCE: 126881309200

; CURRENT APPLICATION NUMBER: US/09/527,487

; CURRENT FILING DATE: 2000-03-16

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 1255

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-527-487-2

Query Match 79.8%; Score 2410.5; DB 4; Length 1255;

Best Local Similarity 41.2%; Pred. No. 1.1e-183;

Matches 510; Conservative 0; Mismatches 2; Indels 727; Gaps 1;

Qy	35	LARGAASQVCTGTDMKRLRPASPEHLDMLRLHLYQGVQVQGNLELYLPTNASLSFLQ	94
Db	16	LPPGAASQVCTGTDMKRLRPASPEHLDMLRLHLYQGVQVQGNLELYLPTNASLSFLQ	75
Qy	95	DIQEVQGYVLIHAHQVQVPLQRLRIVRGTLQFEDNYALAVLDNGDPLNNTTPTVTGASPG	154
Db	76	DIQEVQGYVLIHAHQVQVPLQRLRIVRGTLQFEDNYALAVLDNGDPLNNTTPTVTGASPG	135
Qy	155	GLRELQLRSLTEILKGGVLIQRNPOLCYQDTILWKDIIPHKNQLALTIDITNRSRACHPC	214
Db	136	GLRELQLRSLTEILKGGVLIQRNPOLCYQDTILWKDIIPHKNQLALTIDITNRSRACHPC	195
Qy	215	SPMKSGSCWGESSEDCQSLRTVTCAGGCARCKGPLPTDCCHQCAAGCTGPKHSDCLAC	274
Db	196	SPMKSGSCWGESSEDCQSLRTVTCAGGCARCKGPLPTDCCHQCAAGCTGPKHSDCLAC	255
Qy	275	LHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYVLTSDVGS	329
Db	256	LHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYVLTSDVGSCTLVC	315
Qy	330	-----	329
Db	316	PLHNOEVTAEBDGTQRCCKSPCARVCYGLGMEHLREVRVTSANIQBFAGCKKIFGSLA	375
Qy	330	-----	329
Db	376	FLPESFDGDPASNTAPLQPEQLQVETLEETGYLYISANPDLSLPDLSVFQNLQVIRGRI	435

```

QY 330 ----- 329
Db 436 LHNGAYSLTQIGISWGLSRLSRLGSLALIHNNHLCFVHTVPWDQLFRNPQALLH 495
QY 330 ----- 329
Db 496 TANRPEDECVGEGACHOLCARGHCWPGPTQCVCNSQFLRGQECVEECRVLQGLPREYV 555
QY 330 ----- 329
Db 556 NARHCLPCHPECPQNGSVTCFGEADQCVACAHYKDPFCVACPSGVKPDLSYMPIWK 615
QY 330 ----- 329
Db 616 FPDEGACQPCINCTHSCVDLDDKGCPCAEQASPLTSIVSAVGLLVVLGVVFGILI 675
QY 330 ----- 329
Db 676 KRRQOKIRKYTMRRLQETELVEPLTPSGAMPNQAMRILKETELRKVKVLGSGAGFTVY 735
QY 330 ----- 329
Db 736 KGIWIPDGENVKIPVAIKVLRNTPSKANKKEILDEAYVMAGVSPYVSRLLGICLTSTVQ 795
QY 330 ----- 329
Db 796 LVTQLMPYGCCLLDHVNRGRGLSQDLLNWCQIAKMSYLEDVRLVHRDLAARNVLVKS 855
QY 330 ----- 329
Db 856 PNHVKITDFGLARLLDIDETEHADGGKVPKIKWMALESILRRRFTHQSDVMSYGVTVWEL 915
QY 330 ----- 329
Db 916 MTFGAKPYDGIPAREIPDLLEKGERLPQPICTIDVYIMVKWMIDSECRPRPRELVSE 975
QY 330 ----- 329
Db 976 FSRMARDPQFVVIQNEIDGPASPLDSTFYRSILLEDDMGDLVDAEYLVPOQGFCDP 1035
QY 330 --GAGGMVHRRHSSTRSGGDLTLGLEPSEBEEAPRSLAPSEGAGSDVFDGLGMGAA 387
Db 1036 APGAGGMVHRRHSSTRSGGDLTLGLEPSEBEEAPRSLAPSEGAGSDVFDGLGMGAA 1095
QY 388 KGLQSLPDPSPLORYSEDPTVLPSETDGVVAPLTCSPQPEYVNPQDVRPQPSREG 447
Db 1096 KGLQSLPDPSPLORYSEDPTVLPSETDGVVAPLTCSPQPEYVNPQDVRPQPSREG 1155
QY 448 PLPAARPAGATLERAKTILSPKNGVVKOVFAFGGAVENPEYLTPOGGAAPQHPPPATFSP 507
Db 1156 PLPAARPAGATLERAKTILSPKNGVVKOVFAFGGAVENPEYLTPOGGAAPQHPPPATFSP 1215
QY 508 AFDNLYYWDQPPERGAPESTFKGPTAENPEYLGIDVP 546
Db 1216 AFDNLYYWDQPPERGAPESTFKGPTAENPEYLGIDVP 1254

```

RESULT 2

```

US-09-811-115-3
; Sequence 3, Application US/09811115
; Patent No. 6632979
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A
; CURRENT APPLICATION NUMBER: US/09/811,115
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 3
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-115-3

Query Match      79.8%; Score 2410.5; DB 4; Length 1255;
Best Local Similarity 41.2%; Pred. No. 1.1e-183;
Matches 510; Conservative 0; Mismatches 2; Indels 727; Gaps 1;

QY 35 LARGAASQVCTGDMKRLPASPETHLDMLRHLHYQGCVVQGNLELYLTPTNASLSFLQ 94
Db 16 LPPGAASQVCTGDMKRLPASPETHLDMLRHLHYQGCVVQGNLELYLTPTNASLSFLQ 75
QY 95 DIQEVQGVVLAHQVROVPLQRLIRVGTQOLFEDNYALAVLDNGDPLNNTPPTVGASPG 154
Db 76 DIQEVQGVVLAHQVROVPLQRLIRVGTQOLFEDNYALAVLDNGDPLNNTPPTVGASPG 135
QY 155 GLREIQLRSLTEILKGGVLIQRNPOLCYQDTILMKDIFHKNNQALALTLIDNRSRACHPC 214
Db 136 GLREIQLRSLTEILKGGVLIQRNPOLCYQDTILMKDIFHKNNQALALTLIDNRSRACHPC 195
QY 215 SPMCKGSRCSWGESSEDCOSLTRTVACGACGACRCKPLPTDCCHQCACAGCTGPKHSDCLAC 274
Db 196 SPMCKGSRCSWGESSEDCOSLTRTVACGACGACRCKPLPTDCCHQCACAGCTGPKHSDCLAC 255
QY 275 LHFHNSGICELHCPALVTYNTDTFESMPNPRGYTFGASCVTACPYNYLSTDVGS 329
Db 256 LHFHNSGICELHCPALVTYNTDTFESMPNPRGYTFGASCVTACPYNYLSTDVGSCTLVLC 315
QY 330 ----- 329
Db 316 PLHNQEVTAEDGTQCEKCKPCARVCYGLGMEHLREVRAVTSANIQEFAGCKKIFGSLA 375
QY 330 ----- 329
Db 376 FLPEFSDGDPASNTAPLQEQVLFETLEETGYLYISAWPDSLPLDSVFNQLQVIRGR 435
QY 330 ----- 329
Db 436 LHNGAYSLTQIGISWGLSRLSRLGSLALIHNNHLCFVHTVPWDQLFRNPQALLH 495
QY 330 ----- 329
Db 496 TANRPEDECVGEGACHOLCARGHCWPGPTQCVCNSQFLRGQECVEECRVLQGLPREYV 555
QY 330 ----- 329
Db 556 NARHCLPCHPECPQNGSVTCFGEADQCVACAHYKDPFCVACPSGVKPDLSYMPIWK 615
QY 330 ----- 329
Db 616 FPDEGACQPCINCTHSCVDLDDKGCPCAEQASPLTSIVSAVGLLVVLGVVFGILI 675
QY 330 ----- 329
Db 676 KRRQOKIRKYTMRRLQETELVEPLTPSGAMPNQAMRILKETELRKVKVLGSGAGFTVY 735
QY 330 ----- 329
Db 736 KGIWIPDGENVKIPVAIKVLRNTPSKANKKEILDEAYVMAGVSPYVSRLLGICLTSTVQ 795
QY 330 ----- 329
Db 796 LVTQLMPYGCCLLDHVNRGRGLSQDLLNWCQIAKMSYLEDVRLVHRDLAARNVLVKS 855
QY 330 ----- 329
Db 856 PNHVKITDFGLARLLDIDETEHADGGKVPKIKWMALESILRRRFTHQSDVMSYGVTVWEL 915
QY 330 ----- 329
Db 916 MTFGAKPYDGIPAREIPDLLEKGERLPQPICTIDVYIMVKWMIDSECRPRPRELVSE 975

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:20:15 ; Search time 61.1253 Seconds
(without alignments)
2565.449 Million cell updates/sec

Title: US-09-821-883-1
Perfect score: 3020
Sequence: 1 MRAAPLLARAAASLSGLF.....ENPYGLDVPAAAHHHH 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3020	100.0	555	4	Aae13108 Human HER
2	3005.5	99.5	564	4	Aae13110 Human HER
3	2972	98.4	690	4	Aae13109 Human HER
4	2957.5	97.9	697	4	Aae13111 Human HER
5	2573.5	85.2	919	3	Aab21203 Human HER
6	2573.5	85.2	919	5	Aam51148 Her-2/neu
7	2410.5	79.8	1255	4	Aay92620 Human her
8	2410.5	79.8	1255	4	Aab60167 HER2 tran
9	2410.5	79.8	1255	4	Aae12130 Human tyr
10	2410.5	79.8	1255	5	Aae12130 Human HER
11	2410.5	79.8	1255	5	Aae26349 Human HER
12	2410.5	79.8	1255	5	Aae26366 Human Her
13	2410.5	79.8	1255	5	Aau74545 Human HER
14	2410.5	79.8	1255	6	Abp74708 Breast ca
15	2410.5	79.8	1255	6	Aac38390 Human c-e
16	2410.5	79.8	1255	6	Ada38143 Human erb
17	2410.5	79.8	1255	7	Ada37255 Human Erb
18	2410.5	79.8	1255	7	Adb67621 Human epi
19	2405.5	79.7	1255	2	Aaw01111 HER-2/neu
20	2405.5	79.7	1255	2	Aaw92406 Human HER
21	2405.5	79.7	1255	3	Aay84780 Amino aci
22	2405.5	79.7	1255	3	Aab21198 Human HER
23	2405.5	79.7	1255	4	Aag88267 HER2/neu
24	2405.5	79.7	1255	4	Aab85458 Human HER
25	2405.5	79.7	1255	5	Aae20479 Human HER

26	2405.5	79.7	1255	5	AAU77114	Human Her
27	2405.5	79.7	1255	5	AAU51143	Human Her
28	2405.5	79.7	1255	5	AAE24067	Human Her
29	2405.5	79.7	1255	6	ABR43687	Human C-e
30	2405.5	79.7	1255	7	ABR82066	Human Her
31	2405.5	79.7	1255	7	ADC09593	Her2/Neu
32	2405.5	79.7	1255	7	ADD25484	Binding d
33	2405.5	79.7	1255	7	ADG63281	Human Pro
34	2405.5	79.7	1255	7	ADE76190	Human HER
35	2398.5	79.4	1253	7	ADC35106	Human bre
36	2389.5	79.1	1433	2	AAR39568	Sequence
37	2334.5	77.3	1223	5	AAU98923	Human bre
38	2129	70.5	920	5	AAWS1152	Mouse Her
39	2129	70.5	926	5	AAWS1153	Mouse Her
40	2092.5	69.3	1200	3	AAAB21208	Human HER
41	1961	64.9	1256	3	AAAB21206	Mouse Her
42	1961	64.9	1256	4	AAAG62860	Amino aci
43	1961	64.9	1256	5	AAWS1151	Mouse Her
44	1950.5	64.6	1260	6	ABU79135	ErbB2 (HE
45	1948	64.5	1256	3	AAAB21199	Rat HER-2

ALIGNMENTS

RESULT 1
AAE13108
ID AAE13108 standard; protein; 555 AA.
XX
AC AAE13108;
XX
DT 28-JAN-2002 (first entry)
XX
DE Human HER500 fusion protein construct.
XX
KW Immunostimulatory fusion protein; IFP; antigen component; therapy;
KW Immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human;
KW HER-2 protein; HER500 fusion protein.
XX
OS Homo sapiens.
OS Synthetic.
OS Chimeric.
XX
FN WO200174855-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US010515.
XX
PR 30-MAR-2000; 2000US-0193504P.
XX
(DEND-) DENDREON CORP.
XX
PI Laus R, Vidovic D, Graddie T;
XX
DR WPI; 2001-662965/76.
XX
N-PSDB; AAD21564.

An immunostimulatory fusion protein comprising the intracellular domain of HER-2 and an antigen elicits an immune response to the antigen and is useful for the treatment of associated cancer associated.

Claim 7; Page 26; 59pp; English.

The invention relates to immunostimulatory fusion proteins (IFP) and nucleic acid molecules encoding such proteins. The IFPs comprise a polypeptide antigen component and an immunostimulatory component derived from the intracellular domain of HER-2 protein which is effective to elicit a protective dendritic cell (DC)-induced T-cell mediated cellular immune response to the antigen. IFP or superactivated dendritic cells are

CC used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
 CC associated with a particularly antigen. The present sequence is HER500
 CC fusion protein construct which comprises human PAP signal sequence,
 CC mature PAP protein, an Ala Arg linker, human HER-2 signal sequence,
 CC mature HER-2 membrane distal extracellular and intracellular domains and
 CC a C-terminal tag
 XX
 XX Sequence 555 AA;
 SQ

Query Match 100.0%; Score 3020; DB 4; Length 555;
 Best Local Similarity 100.0%; Pred. No. 6.4e-202;
 Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAAPLLARASISLGLFLFLFWLDRSLAKELARGAASQVCTGDMKRLRPASPET 60
 Db 1 MRAAPLLARASISLGLFLFLFWLDRSLAKELARGAASQVCTGDMKRLRPASPET 60

Qy 61 HLDMLRHLHYQCQVVOGNLELTYLPTNASLSFLQDIOEVQGYVLIANHNRQVPLQRLRI 120
 Db 61 HLDMLRHLHYQCQVVOGNLELTYLPTNASLSFLQDIOEVQGYVLIANHNRQVPLQRLRI 120

Qy 121 VRGTQLFEDNYALAVLDNGDPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOL 180
 Db 121 VRGTQLFEDNYALAVLDNGDPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOL 180

Qy 181 CYQDTILWKDIFHKNNQALTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCA 240
 Db 181 CYQDTILWKDIFHKNNQALTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCA 240

Qy 241 GGCARCKGPLPTDCCHQCAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTPES 300
 Db 241 GGCARCKGPLPTDCCHQCAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTPES 300

Qy 301 MPNPEGYTFGASCVTACPNYLSLTDVSGAGGMVHHRSSSTRSGGDLTLGLEPSEE 360
 Db 301 MPNPEGYTFGASCVTACPNYLSLTDVSGAGGMVHHRSSSTRSGGDLTLGLEPSEE 360

Qy 361 EAPRSLAPSGAGSDVDFDGLGMAAKGLQSLPHTDPSPLQRYSEDPTVPLPSETDGYV 420
 Db 361 EAPRSLAPSGAGSDVDFDGLGMAAKGLQSLPHTDPSPLQRYSEDPTVPLPSETDGYV 420

Qy 421 APLTCSPOPEYVNDVPRPQPPSPREGPLPAARAGATLERAKTSLSPGKNGWKDVFAG 480
 Db 421 APLTCSPOPEYVNDVPRPQPPSPREGPLPAARAGATLERAKTSLSPGKNGWKDVFAG 480

Qy 481 GAVENPEYLTPOGGAAPQHPHPPAPSPAFDNLVYWDQPPRGAPPSTFKTPTAENPEY 540
 Db 481 GAVENPEYLTPOGGAAPQHPHPPAPSPAFDNLVYWDQPPRGAPPSTFKTPTAENPEY 540

Qy 541 LGLDVPAAAHHHHH 555
 Db 541 LGLDVPAAAHHHHH 555

RESULT 2
 AAEL13110
 ID AAEL13110 standard; protein; 564 AA.
 XX
 AC AAEL13110;
 XX
 XX
 DT 28-JAN-2002 (first entry)
 XX
 XX Human HER500 fusion protein construct comprising OVA-derived octapeptide.
 DE
 XX Immunostimulatory fusion protein; IFP; antigen component; therapy;
 KW immunostimulatory component; T-cell mediated immune response; DC;
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
 KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
 KW membrane distal intracellular domain; C-terminal tag; human; OVA;
 KW HER-2 protein; ovalbumin-derived octapeptide; HER500 fusion protein.
 OS Homo sapiens.
 OS Unidentified.

OS Chimeric.
 XX WO200174855-A2.
 XX
 PD 11-OCT-2001.
 XX
 XX 30-MAR-2001; 2001WO-US010515.
 XX
 PR 30-MAR-2000; 2000US-0193504P.
 XX
 PA (DEND-) DENDREON CORP.
 XX
 PI Laus R, Vidovic D, Graddis T;
 XX
 DR WPI; 2001-662965/76.
 DR N-PSDB; AAD21566.
 XX
 PT An immunostimulatory fusion protein comprising the intracellular domain
 of HER-2 and an antigen elicits an immune response to the antigen and is
 useful for the treatment of associated cancer associated.
 PT
 XX
 FS Claim 7; Page 26; 59pp; English.
 XX
 CC The invention relates to immunostimulatory fusion proteins (IFP) and
 nucleic acid molecules encoding such proteins. The IFPs comprise a
 polypeptide antigen component and an immunostimulatory component derived
 from the intracellular domain of HER-2 protein which is effective to
 elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
 immune response to the antigen. IFP or superactivated dendritic cells are
 used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
 associated with a particularly antigen. The present sequence is HER500
 fusion protein construct which comprises human PAP signal sequence,
 mature PAP protein, an Ala Arg linker, human HER-2 signal sequence,
 mature HER-2 membrane distal extracellular domain, an Ala linker, an
 ovalbumin (OVA)-derived immunodominant octapeptide, HER-2 membrane distal
 intracellular domain and a C-terminal tag
 XX
 SQ Sequence 564 AA;
 Query Match 99.5%; Score 3005.5; DB 4; Length 564;
 Best Local Similarity 98.4%; Pred. No. 6.7e-201;
 Matches 555; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

Qy 1 MRAAPLLARASISLGLFLFLFWLDRSLAKELARGAASQVCTGDMKRLRPASPET 60
 Db 1 MRAAPLLARASISLGLFLFLFWLDRSLAKELARGAASQVCTGDMKRLRPASPET 60

Qy 61 HLDMLRHLHYQCQVVOGNLELTYLPTNASLSFLQDIOEVQGYVLIANHNRQVPLQRLRI 120
 Db 61 HLDMLRHLHYQCQVVOGNLELTYLPTNASLSFLQDIOEVQGYVLIANHNRQVPLQRLRI 120

Qy 121 VRGTQLFEDNYALAVLDNGDPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOL 180
 Db 121 VRGTQLFEDNYALAVLDNGDPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOL 180

Qy 181 CYQDTILWKDIFHKNNQALTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCA 240
 Db 181 CYQDTILWKDIFHKNNQALTLIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCA 240

Qy 241 GGCARCKGPLPTDCCHQCAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTPES 300
 Db 241 GGCARCKGPLPTDCCHQCAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTPES 300

Qy 301 MPNPEGYTFGASCVTACPNYLSLTDVSG-----GAGMWHHRSSSTRSGGDL 351
 Db 301 MPNPEGYTFGASCVTACPNYLSLTDVSGASIIINFEKLGAGGMVHHRSSSTRSGGDL 360

Qy 352 TLGLEPSEERAPRSLAPSGAGSDVDFDGLGMAAKGLQSLPHTDPSPLQRYSEDPTVP 411
 Db 361 TLGLEPSEERAPRSLAPSGAGSDVDFDGLGMAAKGLQSLPHTDPSPLQRYSEDPTVP 420

Qy 412 LPSETDGYVAPLTCSPQPEYVNDVPRPQPPSPREGPLPAARAGATLERAKTSLSPGKNG 471
 OS Unidentified.